Universidad Autónoma de Madrid

Optimization Principles in Neurobiology and Collective Animal Behavior

Alfonso Pérez Escudero
Director: Dr. Gonzalo García de Polavieja Embid

Para la obtención del grado de
Doctor
por la Universidad Autónoma de Madrid
Programa de Doctorado en Biofísica
Instituto Nicolás Cabrera
Facultad de Ciencias

Mayo de 2013

Contents

Summary

The aim of this thesis is to understand the selective pressures that have shaped several complex biological systems. Our main tool is the optimization framework. This consists of making a hypothesis about the most relevant selective pressures that have influenced a particular system, finding its optimal configuration according to that hypothesis, and comparing the actual system with the optimal configuration. Here we present applications of this method to the anatomy of the nervous system, to bacterial metabolism and to the rules that govern collective behavior. We also present a tracking method for the study of the behavior of animal groups.

Our first result refers to the anatomy of the nervous system, which appears to be under a great selective pressure to reduce the cost associated with connecting its components to each other. This pressure is described by the principle of wiring economy, which has successfully explained features of the nervous system of many different species. In the nematode *Caenorhabditis elegans*, whose neuroanatomy is known in extraordinary detail, wiring economy gives a good overall description of the nervous system but fails to account for macroscopic features such as the grouping of neurons into ganglia and encephalization. Here we show that the subnetwork of sensory and motor neurons, which comprises of only 15% of connections, is significantly better optimized than the subnetwork of interneurons. Furthermore, we find that the structure of the ganglia is explained by the application of the wiring economy principle on the small subnetwork of sensory and motor neurons.

We also present a general theory of the structure of the deviations from optimality. Using the simple rationale that components with larger impact on total fitness should be closer to their optimal positions than components with a smaller impact, we obtain a testable prediction for systems of many components. Furthermore, the expected pattern of deviations depends not only on the position of the optimum, but also on the shape of the objective function around it. Therefore, in contrast to the usual application of optimization principles, we can distinguish even between objective functions whose optima are at the same point. We find that the deviations of the nervous system of *C. elegans* with respect to its optimal configuration follow the predicted pattern. Furthermore, using the increased discrimination power of our method, we find evidence that wiring cost grows sublinearly with wiring length. Turning to a different system,

we show that deviations of the metabolic network of *Escherichia coli* with respect to the configuration of maximum biomass production follow the expected pattern. While the configuration that maximizes ATP production is almost indistinguishable from that which maximizes biomass, only the deviations corresponding to biomass maximization match the predicted pattern, indicating that it is a more relevant principle than ATP maximization.

We next address the problem of collective behavior, a phenomenon observed in many different species. While there is a great diversity of behavioral mechanisms across species, all of them share the possibility of using the behaviors of other individuals to improve their estimation about the environment. An adequate use of this social information may constitute a significant selective advantage. In order to investigate the impact of this factor on the rules that govern collective behavior, we have developed a model of collective decision-making based on Bayesian estimation of the goodness of each of the available options using both social and private information. We find that, depending on the conditions, the model predicts different ways of taking into account the number of individuals choosing each option. In particular we find three qualitative regimes, that can be associated with different ways of counting. In the first regime the optimal response can be achieved without the need of counting beyond a certain number. The second regime corresponds with using Weber's law to count high numbers, and a different rule to count low numbers. In the third regime it is only necessary to take into account the difference between the numbers of individuals choosing each option, not their absolute numbers. We find a good agreement between the model and experimental data. In particular, the model reproduces very accurately rich datasets of three-spined sticklebacks (*Gasterosteus aculeatus*), argentine ants (*Linepithema humile*) and our own experimental results in zebrafish (*Danio rerio*).

In order to study collective animal behavior experimentally, we have developed a video-tracking technique capable of extracting the trajectory of each animal of a group. Our system is based on a pattern recognition method that allows the identification of each individual in every frame. The method uses the natural physical differences between individuals, not requiring tags of any kind, and works even for size-matched siblings. In contrast to previous techniques that lose the identity of an individual whenever an occlusion cannot be resolved, our system can follow reliably all individuals during arbitrarily long videos.

Resumen

Esta tesis tiene como objetivo estudiar y entender las presiones evolutivas que han moldeado varios sistemas biológicos complejos. Para ello, nos hemos servido principalmente del método de optimización en los sistemas biológicos. Este método consiste en fabricar una hipótesis plausible acerca de las presiones selectivas que pueden haber actuado sobre un sistema, calcular la configuración optima del sistema de acuerdo con dicha hipótesis, y compararamos dicho resultado con la configuración real. Presentamos aplicaciones de este método a la anatomía del sistema nervioso, al metabolismo bacteriano y a las reglas que gobiernan el comportamiento colectivo. Presentamos además el detalle de un sistema de seguimiento para el estudio del comportamiento de animales en grupo.

Nuestro primer resultado se refiere a la anatomía del sistema nervioso, que parece estar afectado por una gran presión selectiva para reducir el coste asociado a la interconexión de sus componentes. Esta presión se describe a través del principio de economía de cable, que ha servido para explicar características de los sistemas nerviosos de muchas especies. Este principio proporciona una buena descripción general del sistema nervioso del nematodo *Caenorhabditis elegans*, cuya neuroanatomía se conoce con extraordinario detalle. Sin embargo, no logra explicar algunas características macroscópicas tales como la agrupación de neuronas en ganglios o la encefalización. En este trabajo mostramos que la subred de neuronas sensoriales y motoras, que comprende tan solo el 15% de las conexiones, se encuentra significativamente mejor optimizada que la subred de interneuronas. Además, encontramos que el principio de economía de cable sí explica la estructura de los ganglios, cuando se aplica exclusivamente sobre la subred de neuronas sensoriales y motoras.

Presentamos también una teoría general sobre la estructura de las desviaciones respecto del óptimo. Nuestro planteamiento se basa en que los componentes que afecten más críticamente a la aptitud (*fitness*) deberían encontrarse más cerca del óptimo que los componentes con menos influencia. De este modo obtenemos predicciones para sistemas de varios componentes que pueden ser comprobadas experimentalmente. Además, encontramos que el patrón de desviaciones respecto del óptimo depende de la forma de la función objetivo alrededor del óptimo, y no sólo de su posición. De este modo podemos incluso distinguir entre funciones objetivo cuyo optimo esta en el mismo punto. Observamos que las desviaciones

del óptimo en el sistema nervioso de *C. elegans* respecto a su configuración óptima siguen el patrón predicho. Además, mediante el uso del mayor poder de discriminación de nuestro método, encontramos evidencias de que el coste de cableado aumenta sublinealmente con la longitud del cable. En un sistema diferente, mostramos que las desviaciones de la red metabólica de *Escherichia coli* con respecto a la configuración que maximiza la producción de biomasa siguen también el patrón esperado. Mientras que la configuración que maximiza la producción de ATP es prácticamente indistinguible de la configuración que maximiza biomasa, tan solo las desviaciones relacionadas con la maximización de biomasa corresponden con el patrón predicho, lo que indica que la maximización de biomasa es un principio más relevante que la maximización de ATP.

A continuación abordamos el problema del comportamiento colectivo, que se da en muchas especies diferentes. Aunque hay una enorme diversidad de mecanismos comportamentales entre diferentes especies, todas tienen en común la posibilidad de usar los comportamientos de los demás individuos para mejorar las estimaciones acerca del entorno. Un uso adecuado de esta información social puede suponer una ventaja evolutiva significativa. Para investigar la influencia de este factor en las reglas que gobiernan el comportamiento colectivo, hemos desarrollado un modelo de toma de decisiones en colectivos de animales basado en la estimación Bayesiana de la calidad de cada una de las opciones mediante el uso de la información social y privada. Encontramos que, según las condiciones, el modelo es capaz de predecir los diferentes modos de tener en cuenta el número de individuos que eligen cada opción. Concretamente encontramos tres regímenes cualitativamente diferentes, que pueden estar asociados con diferentes modos de contar: En el primer régimen la respuesta óptima se alcanza sin la necesidad contar por encima de un cierto número. El segundo régimen corresponde con la ley de Weber para el manejo de grandes números pero no para números pequeños. En el tercer régimen tan solo es necesario tener en cuenta la diferencias entre los números de individuos que eligen cada opción, no los números absolutos. Observamos que el modelo puede reproducir con mucha precisión resultados muy detallados obtenidos por otros investigadores sobre el pez espinosillo (*Gasterosteus aculeatus*) y la hormiga argentina (*Linepithema humile*) y datos de nuestros propios experimentos realizados en el pez cebra (*Danio rerio*).

Para poder observar experimentalmente el comportamiento colectivo de los animales, hemos desarrollado una técnica de seguimiento capaz de extraer de un video las trayectorias de cada individuo del grupo. Nuestro sistema se basa en un método de reconocimiento que permite la identificación de cada individuo en cada fotograma. El método usa las diferencias físicas entre individuos, sin necesidad de marcas artificiales, y funciona incluso en hermanos del mismo tamaño. A diferencia con otras técnicas previas que pierden la identidad del individuo tras una oclusión que no pueda resolverse, nues-

tro sistema permite seguir de manera fiable a todos los individuos durante vídeos de cualquier duración.

1
Introduction

> Owing to this struggle, variations, however slight and from whatever cause proceeding, if they be in any degree profitable to the individuals of a species, *in their infinitely complex relations to other organic beings and to their physical conditions of life*, will tend to the preservation of such individuals, and will generally be inherited by the offspring. [Emphasis added]
>
> Charles Darwin
> *The Origin of Species by Means of Natural Selection*

THIS SENTENCE is the first full enunciation of the principle of natural selection in *The Origin of Species* (Darwin, 1872). While a simpler sentence would have been enough to express the main message, Darwin chose to remark the ability of evolution to blindly solve a problem that very often is intractable for us. Natural selection perceives the adequacy of a given variation regardless of the complexity of the process that makes that variation beneficial or deleterious for its owner.

This stunning contrast between the simplicity of the principles of evolution and the extraordinary complexity of its products is the motivation of this thesis, which is driven by the aim of reducing the intricate details of living beings to relatively simple rules extracted from the basic principles of evolution. This enterprise requires the combination of work in different areas. In the first place, it is essential to have a very good experimental knowledge of the biological systems in question. It is also necessary to understand the process of evolution itself, including the processes that link genotype and phenotype. Finally, it is necessary to know the selective pressures that act on the system. This thesis is focused mainly on the latter of these tasks, building on the work done by many others before us. We therefore consider a few systems that are sufficiently well known to require, if any, only a small experimental effort on our side. For one particular feature of each system, we try to understand the main selective pressures that are responsible for it, and to which extent these selective pressures can explain our observations.

Our basic tool is the optimization framework (Maynard Smith, 1978). This approach consists in essence of the same rationale as we use to solve relatively simple biological questions. For example, why are the optical tissues of the eye transparent? This question can be

answered, at least in a first approximation, simply by knowing that one function of these tissues is to transmit light. Similarly, if we ask why the cornea has a given curvature, one answer is that it is the curvature needed to project, together with the other optical elements of the eye, a sharp image on the retina. These simple answers can be given in those cases in which evolution has produced a system so well adapted to its function that its features are mostly determined by the optimal performance of such function.

The works presented here follow the same simple rationale, only applied to problems in which either the utility of a given feature is not so obvious, or it is not so simple to find out which would be the "ideal" state in which utility would be maximized. Thus, for a given feature of a biological system, we make a hypothesis about how it may contribute to the fitness of the system. We then compute the state of the system that maximizes the fitness according to our hypothesis, and compare the real system with our theoretical result. When theory matches experiment, we have a clue that the hypothesis we made is correct.

In most cases (and in all examples presented in this thesis) this technique only studies the evolution of the phenotype, making the assumption that genetics, development and evolutionary history do not impose very restrictive constraints on the phenotypes that can be realized. While these approximations will not be valid in many cases, this approach has led to useful insights in many systems, for example the genetic code (Freeland and Hurst, 1998), expression level of proteins (Dekel and Alon, 2005), bacterial metabolism (Varma and Palsson, 1994), the anatomy of the vascular system (Murray, 1926), neuroanatomy (Ramón y Cajal, 1899; Chen et al., 2006), behavior (Oaten, 1977; Parker and Simmons, 1994) and many others. Even in the cases in which the approximations do not hold, and important features of the system are due, for example, to the evolutionary history of the species, the importance of the other factors is only recognized after realizing that the system is far from its optimal configuration.

The first question that we ask in this thesis is about the anatomy of the nervous system. In spite of its great complexity, a great part of its features can be explained by a single principle, known as the wiring economy principle, which was formulated by Ramón y Cajal (1899).

> Pero en cuanto hubimos descartado la necesidad del paso de la conmoción nerviosa por el soma, todo fueron facilidades; pues caímos en la cuenta de que las referidas dislocaciones constituyen adaptaciones morfológicas regidas por la ley de economía (leyes de ahorro de tiempo, de espacio y de materia).[1]
>
> Santiago Ramón y Cajal
> *Textura del Sistema Nervioso del Hombre y de los Vertebrados*

[1] As soon as we rejected the need for the passage of the nervous impulse through the cell body, everything became easy; we realized that the mentioned dislocations constitute morphological adaptations governed by the law of economy (laws of saving of time, space and matter)

In more general terms, this principle states that the morphology of the nervous system will be such that it minimizes the amount of wiring needed to connect its components. This principle has been

tested in many situations, and recently it was found to hold approximately in the nervous system of *Caenorhabditis elegans* (Chen et al., 2006). Thanks to an extremely detailed and complete study of the neuroanatomy of this nematode, the positions of most of its neurons and the connectivity among them are known. Chen et al. (2006) were thus capable of testing the wiring economy principle in a whole organism at the level of individual neurons. They found a good overall agreement between the anatomy of the system and the predictions of the wiring economy principle, but with some significant discrepancies.

Our study of this system focused not on the optimum itself, but rather on the deviations from that optimum. Our fist finding was that the subnetwork of sensory and motor neurons is significantly nearer to the predicted optimum than the network of interneurons. More importantly, it had been found that the macroscopic anatomical features of the nervous system (in particular, the spatial segregation of neurons in 10 ganglia) could not be explained by the principle of wiring economy (Chen et al., 2006). We found that this macroscopic structure arises naturally from the principle of wiring economy, applied only on the subnetwork of sensory and motor neurons, even though this subnetwork only contains 15% of the connections of the total system. These results are presented in CHAPTER 2.

But probably our most interesting discovery on this system was a general observation about the deviations from the optimum. One important limitation of optimization principles is that even in an ideal situation in which we knew the exact form of the selective pressure acting on our system, and in which genetics and development imposed no constraints, we should not expect biological systems to be perfectly optimal. Evolution is a stochastic process, and sometimes variations that are neutral or even deleterious survive, while beneficial ones become extinct.

Then, what deviations can we attribute to the stochasticity of evolution, and which ones are a hint that we are missing something in our understanding of the system? This question has a great practical importance, because one of the main uses of optimization studies is to find out whether our hypotheses about the function of the system are correct and complete. Therefore it is our duty to investigate if alternative hypotheses achieve a better fit to the data. But it is always possible to add complexity to a model and improve a fit, and it is easy to go too far trying to explain deviations that are simply the natural result of evolution. This difficulty to gauge the deviations from optimality has been one of the reasons why the use of the optimization framework in biology has been strongly criticized (Gould and Lewontin, 1979; Parker and Maynard Smith, 1990).

The nervous system of *C. elegans* led us to a way to distinguish deviations that result from the stochastic nature of evolution. While deviations from optimality are to be expected, it is reasonable to think that deviations with a big impact on fitness will be less likely than deviations with a smaller impact. Therefore, in a system of many

components we expect that those components that have a greater impact on fitness will be nearer to their optimal configuration than components with a smaller impact. This simple idea provides us with a testable hypothesis: There must be a correlation between the impact of each component on fitness and their deviations. This pattern is a signature of the evolutionary process, and in general would not be present if deviations were due to a mistake in the hypotheses of the model. Furthermore, because this pattern depends on the shape of the objective function around the optimum, it contains additional information about it. While the usual optimization approach can only distinguish objective functions whose optima are at different points, using the pattern of deviations we can distinguish even between functions whose optima are the same. We found that the deviations of the nervous system of *C. elegans* follow the predicted pattern, with less connected neurons more deviated on average than more connected ones. Also, many different functional forms for the dependency of cost on wire length provide optimal configurations compatible with the experimental data. Using the deviations from the optimum we were able to distinguish them, finding that a sublinear increase of wiring cost with wire length is significantly more compatible with the data than alternative hypotheses.

In order to prove the generality of this result, we turned to a completely different system: the metabolism of the bacterium *Escherichia coli*. This metabolism consists of a network of chemical reactions that, thanks to the amazing work of thousands of researchers, are known with extreme detail. It is reasonable to think that this network is optimized to produce the maximum growth for the bacteria, and it is possible to compute the state of the network that produce this maximum output (Varma and Palsson, 1994). We found again in this system that deviations from the optimum were larger in those components with smaller impact on fitness. Furthermore, while previous studies found that both maximization of biomass production and maximization of ATP were equally compatible with the experimental results (Schuetz et al., 2007), we found that deviations are significantly more compatible with maximization of biomass production. Our theory on the deviations from optimality and the results in *C. elegans* and *E. coli* are presented in CHAPTER 3.

The last system that we studied is the collective behavior of social animals. This field has a long and successful tradition of quantitative analysis, based on extracting simple individual rules from which the collective patterns arise (Goss et al., 1989; Couzin, 2009). Less attention has been paid to the evolutionary origin of such rules, that in most cases have been obtained heuristically. Our aim was to build a theory of collective decision-making based on a relevant and general selective pressure. A very general determinant of behavior is the need to make correct estimations about the environment with limited and ambiguous information. Using a Bayesian framework, we built a theory in which each individual uses the behaviors of other individuals of the group to improve its estimations about the environment

and make better decisions. This same rationale is used in the theory of information cascades, widely used by economists (Banerjee, 1992; Bikhchandani et al., 1992). Our theory is similar, but there are key differences in the way in which the estimation is used to make the final decision. Due to these differences, we make quantitative probabilistic predictions about the decision-making behavior of social animals that sometimes differ from those of information cascades. We find that the predictions of our model correspond accurately with experimental results in several species. While in general many other factors besides the use of social information will influence collective behavior, the good results of our model suggests that information may be responsible of an important share of the selective pressure acting on the system. CHAPTER 4 presents a first version of the theory, and its application to a rich dataset of two-choice decisions in three-spined stickleback (*Gasterosteus aculeatus*) (Ward et al., 2008; Sumpter et al., 2008). CHAPTER 5 presents a more general version of the theory, that accurately explains, in addition to the stickleback dataset, another very detailed dataset of the behavior of argentine ants (*Linepithema humile*) (Perna et al., 2012) and our own experimental results in zebrafish (*Danio rerio*).

In order to study experimentally the collective behavior of animals, we needed a system to extract the trajectory of each individual of a group. There are powerful tracking systems for the study of animal behavior (Branson et al., 2009; de Chaumont et al., 2012; Delcourt et al., 2012; Butail and Paley, 2012), but none of them is capable of automatically maintaining correct identities of individuals during long videos. We have developed a new tracking system that automatically extracts from a video the trajectory of each individual of a group. Our system is based on the identification of each individual in each frame based on the natural physical differences between individuals. It is valid in principle for any species, and can track reliably videos of any duration. This system was intended to serve as Methods section of Chapters 4 and 5 but, not surprisingly, things did not go as planned. Our theoretical and experimental work advanced faster than the development of the tracking system, and our results in Chapters 4 and 5 do not rely on it. However, we consider the tracking system itself as a valuable contribution, and therefore we describe it in CHAPTER 6.

Chapters 2-5 present work that has been already published, and therefore consist of a reproduction of the published papers (Pérez-Escudero and de Polavieja, 2007; Pérez-Escudero et al., 2009; Pérez-Escudero and de Polavieja, 2011; Arganda et al., 2012). The contents of Chapter 6 are currently being prepared for peer-reviewed publication.

Bibliography

Arganda, S., Pérez-Escudero, A., and de Polavieja, G. G. (2012). A common rule for decision-making in animal collec-

tives across species. *Proceedings of the National Academy of Sciences*, 109(50):20508–20513.

Banerjee, A. V. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, 107(3):797–817.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100(5):992–1026.

Branson, K., Robie, A. A., Bender, J., Perona, P., and MH, D. (2009). High-throughput ethomics in large groups of Drosophila. *Nature Methods*, 6(6):451–457.

Butail, S. and Paley, D. A. (2012). Three-dimensional reconstruction of the fast-start swimming kinematics of densely schooling fish. *Journal of the Royal Society, Interface*, 9(66):77–88.

Chen, B. L., Hall, D. H., and Chklovskii, D. B. (2006). Wiring optimization can relate neuronal structure and function. *Proceedings of the National Academy of Sciences of the United States of America*, 103(12):4723–4728.

Couzin, I. D. (2009). Collective cognition in animal groups. *Trends in cognitive sciences*, 13(1):36–43.

Darwin, C. (1872). *The Origin of Species by Means of Natural Selection; or The Preservation of Favoured Races in the Struggle for Life*. Project Gutenberg. Web. November 23, 2009 (ebook #2009), 6th edition.

de Chaumont, F., Coura, R. D.-S., Serreau, P., Cressant, A., Chabout, J., Granon, S., and Olivo-Marin, J.-C. (2012). Computerized video analysis of social interactions in mice. *Nature methods*, 9(4):410–417.

Dekel, E. and Alon, U. (2005). Optimality and evolutionary tuning of the expression level of a protein. *Nature*, 436(7050):588–592.

Delcourt, J., Denoël, M., Ylieff, M., and Poncin, P. (2012). Video multitracking of fish behaviour: a synthesis and future perspectives. *Fish and Fisheries*, 14(2):186–204.

Freeland, S. J. and Hurst, L. D. (1998). The genetic code is one in a million. *Journal of molecular evolution*, 47(3):238–248.

Goss, S., Aron, S., Deneubourg, J., and Pasteels, J. (1989). Self-organized shortcuts in the Argentine ant. *Naturwissenschaften*, 76(12):579–581.

Gould, S. J. and Lewontin, R. C. (1979). The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B: Biological Sciences*, 205(1161):581–598.

Maynard Smith, J. (1978). Optimization theory in evolution. *Annual Review of Ecology and Systematics*, 9:31–56.

Murray, C. D. (1926). The physiological principle of minimum work: I. The vascular system and the cost of blood volume. *Proceedings of the National Academy of Sciences of the United States*, 12(3):207–214.

Oaten, A. (1977). Optimal foraging in patches: a case for stochasticity. *Theoretical population biology*, 12(12):263–285.

Parker, G. and Maynard Smith, J. (1990). Optimality theory in evolutionary biology. *Nature*, 348:27–33.

Parker, G. A. and Simmons, L. W. (1994). Evolution of phenotypic optima and copula duration in dungflies. *Nature*, 370(6484):53–56.

Pérez-Escudero, A. and de Polavieja, G. G. (2007). Optimally wired subnetwork determines neuroanatomy of Caenorhabditis elegans. *Proceedings of the National Academy of Sciences of the United States of America*, 104(43):17180–17185.

Pérez-Escudero, A. and de Polavieja, G. G. (2011). Collective animal behavior from Bayesian estimation and probability matching. *PLoS computational biology*, 7(11):e1002282.

Pérez-Escudero, A., Rivera-Alba, M., and De Polavieja, G. G. (2009). Structure of deviations from optimality in biological systems. *Proceedings of the National Academy of Sciences of the United States of America*, 106(48):20544–20549.

Perna, A., Granovskiy, B., Garnier, S., Nicolis, S. C., Labédan, M., Theraulaz, G., Fourcassié, V., and Sumpter, D. J. T. (2012). Individual rules for trail pattern formation in Argentine ants (Linepithema humile). *PLoS computational biology*, 8(7):e1002592.

Ramón y Cajal, S. (1899). *Textura del sistema nervioso del hombre y de los vertebrados*. Reprinted by Gobierno de Aragón, 2002.

Schuetz, R., Kuepfer, L., and Sauer, U. (2007). Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Molecular systems biology*, 3:119.

Sumpter, D. J. T., Krause, J., James, R., Couzin, I. D., and Ward, A. J. W. (2008). Consensus decision making by fish. *Current Biology*, 18(22):1773–1777.

Varma, A. and Palsson, B. O. (1994). Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type Escherichia coli W3110. *Applied and environmental microbiology*, 60(10):3724–3731.

Ward, A. J. W., Sumpter, D. J. T., Couzin, I. D., Hart, P. J. B., and Krause, J. (2008). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19):6948–6953.

2
An optimal subnetwork in the nervous system of C. elegans

Optimally wired subnetwork determines neuroanatomy of *Caenorhabditis elegans*

Alfonso Pérez-Escudero* and Gonzalo G. de Polavieja†

Neural Processing Laboratory, Instituto ''Nicolás Cabrera'' de Física de Materiales, and Department of Theoretical Physics, Universidad Autónoma de Madrid, 28049 Madrid, Spain

Edited by Solomon H. Snyder, Johns Hopkins University School of Medicine, Baltimore, MD, and approved September 7, 2007 (received for review April 5, 2007)

Wiring cost minimization has successfully explained many structures of nervous systems. However, in the nematode *Caenorhabditis elegans*, for which anatomical data are most detailed, wiring economy is thought to play only a partial role and alone has failed to account for the grouping of neurons into ganglia [Chen BL, Hall DH, Chklovskii DB (2006) *Proc Natl Acad Sci USA* 103:4723–4728; Kaiser M, Hilgetag CC (2006) *PLoS Comput Biol* 2:e95; Ahn Y-Y, Jeong H, Kim BJ (2006) *Physica A* 367:531–537]. Here, we test the hypothesis that optimally wired subnetworks can exist within nonoptimal networks, thus allowing wiring economy to give an improved prediction of spatial structure. We show in *C. elegans* that the small subnetwork of wires connecting sensory and motor neurons with sensors and muscles, comprising only 15% of connections, is close to optimal and alone predicts the main features of the spatial segregation of neurons into ganglia and encephalization. Moreover, a method to dissect networks into optimal and nonoptimal components is shown to find a large near-optimal subnetwork of 84% of neurons with a very low position error of 5.4%, and that explains clustering of neurons into ganglia and encephalization to fine detail. In general, we expect realistic networks not to be globally optimal in wire cost. We thus propose the strategy of using near-optimal subnetworks to understand neuroanatomical structure.

encephalization | ganglia | optimization | wiring economy | anatomy

Why do neurons organize spatially in ganglia, strata, brain areas, and maps? Spatial patterns might have several origins. Evolutionary and developmental constraints underlie spatial organization in nervous systems (1). Neuromodulation can be more effective when neurons form spatial groups (2). Processing can require the use of different wire lengths and a spatial organization of delays (3). Wiring economy (4) can also explain spatial patterns and has the advantage of being quantitatively testable. This principle states that the morphology of nervous systems is such that the cost of interconnecting its parts is the minimum possible given that the conditions for the proper functioning of the system are met. It is simple to see that wiring economy can result in the organization of the nervous system into spatial structures. Imagine, for example, the simple case of two types of neurons, say, R or L neurons responding to the excitation of right and left eyes in a human, respectively. If R neurons are heavily connected among themselves and the same for the L neurons, whereas the R–L connections are few, it can save wire to have the R neurons and L neurons segregated into two different spatial groups, because mixing them would increase the amount of wire used.

The wiring economy principle has been applied to many structures in nervous systems. It has been used to explain large brain structures, such as the existence of brain areas (5), neocortical folding (6), retinotopic maps (7, 8), ocular dominance patterns (5, 9), orientation maps (5, 10, 11), segregation of gray and white matter (12–14), and the arrangement of areas in vertebrate cortex and ganglia in *C. elegans* (15–18). Wiring economy has also been used to explain the structure of neurons, including their dimensions (19–21) and branching angles (22, 23).

The most detailed application of wiring economy has been to the entire nervous system of *Caenorhabditis elegans* (24). These authors updated the wiring diagram of *C. elegans* (25, 26) and used methods developed previously for circuit board design (27, 28) to predict neuronal layout (18). Predicted positions differed from actual ones on average $<10\%$ the length of the animal. These results showed that wiring economy is at work in *C. elegans*, but the layout was found to be nonoptimal (29, 30). The actual spatial organization in distinct ganglia could not be obtained from wiring economy applied to the complete network unless extra factors related to axonal guidance and command neuron function were added to the theoretical prediction (24).

Here, we use wiring economy alone to predict the clustering of neurons in ganglia in *C. elegans*. The article is organized as follows. First, we show that the optimization of the subnetwork of connections from sensory and motor neurons to sensors and muscles, respectively, is responsible for the main features of the clustering in ganglia. Second, we show that wiring economy further applied to the connections among sensory and motor neurons lowers the mean error of predicted neuron positions and improves details of the prediction of clustering. Third, we propose a method to dissect networks into optimal and nonoptimal components. Application to the experimental configuration of neurons in *C. elegans* finds an optimal subnetwork of 84% of neurons that explains the neuroanatomy of the animal down to fine details and using only wiring economy.

Results

Wiring Optimization of Complete Network Does Not Predict Actual Clustering. The *C. elegans* network is formed by 279 neurons, excluding pharyngeal and unconnected neurons, organized in clusters known as ganglia (Fig. 1*A*; see ref. 25). There are 10 ganglia known as anterior (G1 in Fig. 1*A*), dorsal (G2), lateral (G3), ventral (G4), retrovesicular (G5), posterolateral (G6), ventral cord (G7), preanal (G8), dorsorectal (G9), and lumbar (G10). Ganglia G1–G5 are clustered in the head and G8–G10 in the tail. The head also has a high concentration of synapses in a region known as nerve ring. There are 73 sensory neurons (neurons connected to sensors and other neurons), 113 motor neurons (neurons connected to muscles and other neurons), 13 neurons that are simultaneously sensory and

Author contributions: A.P.-E. and G.G.d.P. designed research, performed research, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

*Present address: Instituto de Óptica ''Daza de Valdés,'' Consejo Superior de Investigaciones Científicas, Serrano 121, 28006 Madrid, Spain.

†To whom correspondence should be addressed. E-mail: gonzalo.polavieja@uam.es.

This article contains supporting information online at www.pnas.org/cgi/content/full/0703183104/DC1.

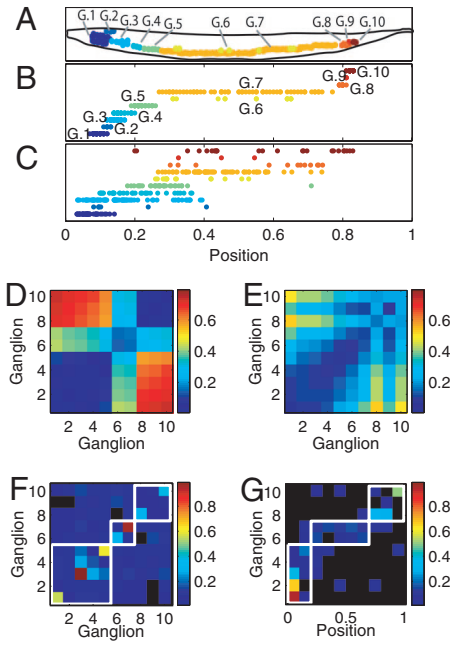


Fig. 1. Optimization of the complete *C. elegans* network fails to predict actual clustering. (*A*) Soma positions in the actual nematode are organized in spatial clusters known as ganglia. (*B*) Soma positions in the actual nematode. Somas in different ganglia are offset for clarity. (*C*) Soma position obtained by minimizing the wiring cost of the complete network. Neurons in different ganglia in the actual network are offset for clarity. (*D*) Average distance between somas belonging to the same ganglion (diagonal elements) and different ganglia (nondiagonal elements) for the actual nematode. (*E*) Same as *D* but obtained for soma positions minimizing the total wiring cost. (*F*) Average number of connections per neuron between ganglia. Diagonal squares, connections between neurons belonging to the same ganglion; nondiagonal squares, connections between neurons belonging to different ganglia. White boxes enclose regions of the graph corresponding to ganglia overlapping in the one dimension considered. Elements with a value of exactly 0 are colored in black. (*G*) Average number of connections per neuron between ganglia and organs that lie in each of the 10 bins we divided the length of the nematode. White boxes enclose regions corresponding to connections between ganglia and organs located at the same position in the body as the ganglion. Elements with a value of exactly 0 are colored in black. *C* and *E* were obtained with $\alpha = \beta = 1/29.3$.

motor, and 80 interneurons (neurons connected only to other neurons).

We adopt the "dedicated wire" cost model of (24) for the *C. elegans* network. Because the length of this nervous system is more than 10 times its diameter, the model considers a single dimension for simplicity. The network is modeled as a set of nodes representing the cell bodies connected by wires, which represent axons and dendrites. Although in *C. elegans* each axon makes more than one synapse, the model in principle assumes that each synapse has a dedicated wire. The total wiring cost W is written as the sum of three costs, one for connections among neurons, a second one for connections between neurons and sensors, and a third one for connections between neurons and muscles

$$W = \frac{1}{2}\alpha \sum_{i,j=1}^{N} A_{ij}|x_i - x_j|^{\xi} + \sum_{i,k=1}^{N,S} B_{ik}|x_i - s_k|^{\xi} + \beta \sum_{i,l=1}^{N,M} C_{il}|x_i - m_l|^{\xi}, \qquad [1]$$

with A_{ij}, B_{ik}, and C_{il} the number of synapses between neurons i and j, neuron i and sensor k, and neuron i and muscle l, respectively. N, S, and M are the number of neurons, sensors, and muscles, respectively. x_i is the 1D projection of the position of neuron i, with s_k and m_l analogously for the fixed positions of sensor k and muscle l, respectively. All positions have values from 0 to 1 (where 0 is the head and 1 is the tail). The 1/2 factor multiplying A_{ij} takes into account the fact that the distance between each pair of neurons is summed twice (note that A is a symmetric matrix). α and β are normalization constants to take into account differences in average cost of the sensory neuron, motor neuron, and interneuron groups. These coefficients can account, for example, for part of the difference between the structure of the "dedicated wire" model and the real network. In the actual network, each neurite that connects two neurons or a neuron and a muscle holds on average 29.3 synapses. Because the model assumes a wire per synapse, the extra cost can be discounted by dividing the wire cost of neuron-to-neuron and neuron-to-muscle connections by 29.3, that is, by making $1/\alpha = 1/\beta = 29.3$ (24). On the other hand, actual connections between neurons and sensors do have a dedicated neurite for each synapse, so, in this case, the model is already in good agreement with reality. Wire cost may be due, for example, to wire volume (5, 18, 22), signal delay and attenuation (31, 32), metabolic costs (33, 34), noise and developmental costs (35). Its functional form is unknown but it is clear that it should increase with wire length. Here, it is written as a power ξ of the wire length. Because wiring cost minimization is analytical for the quadratic case, $\xi = 2$, results have always been obtained for this exponent except where indicated otherwise. Numerical tests suggest that these results are robust with respect to a change in the precise form of the cost (see *Materials and Methods*). The optimal neuron positions minimizing the total cost W in Eq. **1** for $\xi = 2$ can be obtained analytically as follows [see refs. 18, 27, and 28 and supporting information (SI) *Text*]:

$$\vec{x} = Q^{-1}[B\vec{s} + \beta C\vec{m}] \qquad [2a]$$

$$Q_{ip} = \delta_{ij}\left(\alpha \sum_{j=1}^{N} A_{ij} + \sum_{k=1}^{S} B_{ik} + \beta \sum_{l=1}^{M} C_{il}\right) - \alpha A_{ip}. \qquad [2b]$$

The optimal positions of the 279 nonpharyngeal neurons calculated from Eq. **2** and using the connectivity matrices and sensor and muscle positions of the real nematode (see *Materials and Methods*) compare well with the actual positions of the somas with a low average error of $e_p = 9.71\%$, much lower than the average error of random positions, $e_{p,rand} = 34.6\%$ (24). However, the correct clustering is not predicted by wire cost minimization of the full network (Fig. 1 *B–E*). Neurons within each ganglion are closer to each other than to neurons in other ganglia, although in the 1D projection some ganglia overlap. The first five ganglia, located in the head of the nematode, are compact and overlap in the 1D projection. The same happens with the last three, which are located in the tail. The other two (ganglia 6 and 7) are much more elongated and located at the mid-body region. Fig. 1*B* shows the positions of the neurons along the actual nematode, grouped by ganglia, and offset for clarity. Fig. 1*C* shows the positions of the predicted positions of somas obtained from wire cost minimization applied to the complete network. As in Fig. 1*B*, neurons are grouped by ganglia in the actual animal and offset for clarity. Neurons belonging to the same ganglion in the actual network are now dispersed. Fig. 1*D* shows for the actual nematode the mean distance between neurons in the same ganglion (diagonal squares), and between neurons that lie in different ganglia (off-diagonal squares) (see *Materials and Methods*). The clustering structure seen in the actual nematode, Fig. 1*D*, is however very different to the one predicted by wire cost minimization applied to the complete network (Fig. 1*E*; see also Fig. 1*C* for an intuitive picture). To quantify how far the predicted clustering structure is from the actual one, we subtracted the clustering graphs in Fig. 1 *D* and *E* and calculated the mean resulting value as the clustering error, $e_c = \Sigma_{ij}^{G}|a_{ij}(\text{actual}) - a_{ij}(\text{predicted})|/G^2$, where

NEUROSCIENCE

$G = 10$ is the number of ganglia, and a_{ij}(actual) and a_{ij}(predicted) are the elements of the actual and predicted clustering graphs, respectively. The clustering error e_c has values between 0 (perfect prediction) and 1 (animal length), and we multiply it by 100 to give values as percentage of animal length, similarly to the neuron position error. Its value for the optimization of the complete network is $e_c = 16.4\%$ the length of the animal, a value closer to the case of random positions, $e_{c,rand} = 21.8\%$, than to a perfect prediction. In the following, we always use the notation e_c for clustering error as defined above and e_p for average position error.

Although the actual clustering in *C. elegans* is not consistent with wiring cost minimization of the complete network, we wondered whether it could be consistent with wiring economy of subnetworks. If clustering is due to wiring economy, there could be in principle the following two simple scenarios for its origin. Wiring economy would predict clustering when there are more connections within a group of neurons than between neurons belonging to different groups. In this case, neurons belonging to the same group or ganglion would cluster to minimize the cost of intra-ganglion wires. Fig. 1*F* shows the number of connections per neuron between the neurons of the same ganglion (diagonal elements) and between neurons of different ganglia (off-diagonal elements). Elements representing exactly 0 connections per neuron are colored in black. Clustering would naturally arise in the optimal layout if diagonal elements were much stronger than nondiagonal ones. Furthermore there should be much higher connectivity between ganglia that are close to each other (e.g., ganglia 1 to 5) than between distant ganglia (e.g., ganglia 3 and 10). White lines in Fig. 1*F* limit these regions of the graph for which high connectivity between ganglia would be consistent with the actual layout. For example, elements within the big white box in the lower-left corner represent connectivity among the first five ganglia, which are very close to each other in the actual nematode. The other two white boxes have a similar interpretation. Elements outside the white boxes represent connectivity between distant ganglia. Although elements with highest connectivity values lie within the white boxes, elements outside them are also important with each ganglion connected to almost all ganglia, so a simple picture for clustering does not emerge.

An alternative explanation for the clustering of neurons in ganglia would consist in a high interconnectivity between neurons and sensors and muscles. Fig. 1*G* shows the number of connections per neuron between each ganglion and the organs, sensors and muscles, located within small segments of the nematode. In this case, the connectivity is much more consistent with the actual layout of the ganglia: ganglia 1 to 5 are strongly connected to a small region at the head of the nematode, ganglia 6 and 7 are connected to a large region in the middle of the body, and ganglia 8 to 10 are connected to a small region at the tail. Vertical white lines in Fig. 1*G* limit these regions of connectivity, and horizontal white lines separate the three groups of ganglia. It is clear that each ganglion is strongly connected only to one region of the animal, and very weakly connected, or not connected at all (black elements in Fig. 1*G*), to organs in other regions. This connectivity pattern suggests a simple picture for clustering based on wiring economy and a high interconnectivity between sensory and motor neurons and spatial patches of organs.

Wiring Optimization of Subnetwork of Connections from Sensory and Motor Neurons to Organs Predicts Main Features of Clustering. We therefore tested the prediction of cost minimization of wires connecting sensory and motor neurons with sensors and muscles, respectively. The associated cost for this case can be obtained from the total cost in Eq. **1** making $\alpha = 0$, that is, eliminating from the optimization the costs of neuron-to-neuron wires,

$$W = \sum_{i,k=1}^{\hat{N},S} B_{ik}|x_i - S_k|^{\xi} + \beta \sum_{i,l=1}^{\hat{N},M} C_{il}|x_i - m_l|^{\xi}, \qquad [3]$$

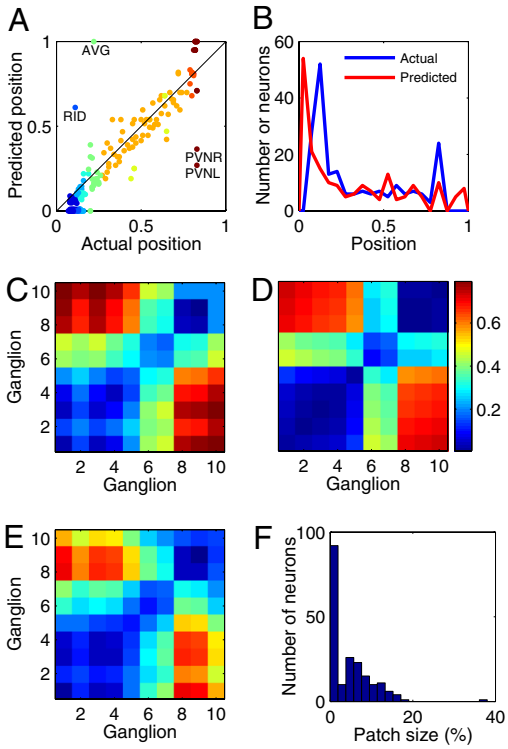


Fig. 2. Optimization of neuron-to-organ connections predicts main features of clustering. (*A*) Optimal positions of sensory and motor neurons, obtained by minimizing the cost of wiring them to organs, versus their actual positions. Colors distinguish different ganglia. (*B*) Density sensory and motor neurons along the nematode for actual and optimal positions. (*C*) Average distance between neurons belonging to the same ganglion (diagonal elements) or different ganglia (nondiagonal elements) for optimized sensory and motor neuron positions. (*D*) Same as *C* but for the actual nematode. (*E*) Same as *C* but optimizing the complete network. (*F*) Histogram of sizes of the minipatches of skin to which each neuron connects.

with $\hat{N} = 199$ the total number of sensory and motor neurons. Minimization of this cost function for cost exponent $\xi = 2$ gives explicitly the optimal position for sensory and motor neurons of the form (see *SI Text*)

$$x_i = \frac{\sum_{k=1}^{S} B_{ik} s_k + \beta \sum_{l=1}^{M} C_{il} m_l}{\sum_{k=1}^{S} B_{ik} + \beta \sum_{l=1}^{M} C_{il}}. \qquad [4]$$

Therefore, wiring economy in this case reduces to a simple prediction by which each sensory and motor neuron should be located at the center of mass of the sensors and muscles it connects to. Despite its simplicity, the center of mass formula (Eq. **4**) predicts to a very good approximation the neuronal layout of the 199 sensory and motor neurons of *C. elegans* (mean position error $e_p = 9.08\%$, $r = 0.923$ for $\beta = 1$) (Fig. 2*A*). Also, the density of the neurons along the animal is similar to the actual one, predicting encephalization as seen in the real animal although with brain ganglia shifted to a more anterior position (Fig. 2*B*). More importantly, the main features of the clustering pattern in ganglia are correctly predicted. The predicted distances between neurons from the same ganglion and between neurons from different ganglia (Fig. 2*C*, clustering error $e_c = 6.99\%$), compare well with the actual clustering pattern of sensory and motor neurons (Fig. 2*D*).

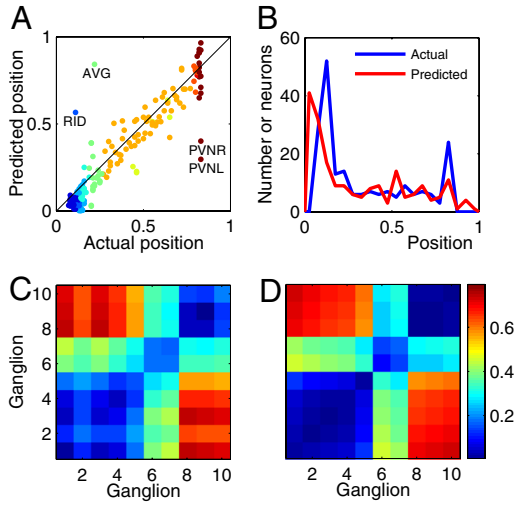


Fig. 3. Connections among sensory and motor neurons improve predicted clusters. (*A*) Optimal positions of sensory and motor neurons, obtained by minimizing their connections to organs and among themselves, versus their actual positions. Colors distinguish different ganglia. (*B*) Neuron density along the nematode, for actual and optimal positions. (*C*) Average distance between neurons belonging to the same ganglion (diagonal elements) or different ganglia (nondiagonal elements) for optimized sensory and motor neuron positions. (*D*) Same as *C* but for the actual nematode.

The predictions obtained using the center of mass formula in Eq. **4** are extremely robust. Varying the parameters β and the cost power ξ have an effect on average position error of 1% and on clustering error <0.3% because of the fact that sensors and muscles are connected to small skin patches (Fig. 2*F*; see also *SI Text*). It is interesting to compare the results of this simple center-of-mass calculation with the predictions for sensory and motor neurons when optimizing the complete network, Eq. **2**. The difference between the two calculations is that the complete network one also includes all neuron-to-neuron connections. For the complete network calculation we found that average position error e_p and clustering error e_c are lowest for $\alpha = 0.05$ and $\beta = 1.5$. However, even for this best case the clustering is worse than that obtained using the simple center-of-mass calculation (Fig. 2*E*, $e_c = 7.26\%$). For other values like $\alpha = \beta = 1/29.3$ chosen in reference (24) the clustering error is even higher (SI Fig. 6*A*, $e_c = 12.3\%$). We thus conclude that the net effect of the neuron-to-neuron connections in the optimization worsens the predictions. The rest of the paper is dedicated to prove that predictions can however be improved further by including part of the neuron-to-neuron connections. In this way we will be able to predict neuron positions for more than just the 199 sensory and motor neurons and explain neuroanatomy to finer details.

Optimal Connections Among Sensory and Motor Neurons Improve Predicted Clusters. We started by adding the connections among sensory neurons, motor neurons and between sensory and motor neurons. The effect of these neuron to neuron connections is obtained from wiring economy of the subnetwork of sensory and motor neurons using the total cost W in Eq. **1** but restricting the sums to sensory and motor neurons. We found an improvement over the center of mass formula in Eq. **4** for neuronal positions, clustering structure and neuron statistics along the animal for a large parameter range, $\alpha \leq 0.1$ and $\beta \geq 0.06$. The best results for neuronal positions and clustering structure were found for $\alpha = 0.05$ and $\beta = 1.5$ (in the remaining, we always use these parameter values). Neuronal layout (Fig. 3*A*, $e_p = 7.71\%$, $r = 0.93$) and clustering in ganglia (Fig. 3 *B* and *C*; $e_c = 5.23\%$) are closer to actual ones (Fig. 3*D*).

Dissection into Optimal and Nonoptimal Subnetworks. To proceed further, we needed a method to dissect the complete network into optimal and nonoptimal subnetworks in the sense of Eq. **2**. The number of subnetworks in a system of N neurons is 2^N, an astronomically large number for analysis. Instead, we have used a method of analysis that uses of the order of N calculations. In the following, we illustrate the core of our method using the toy

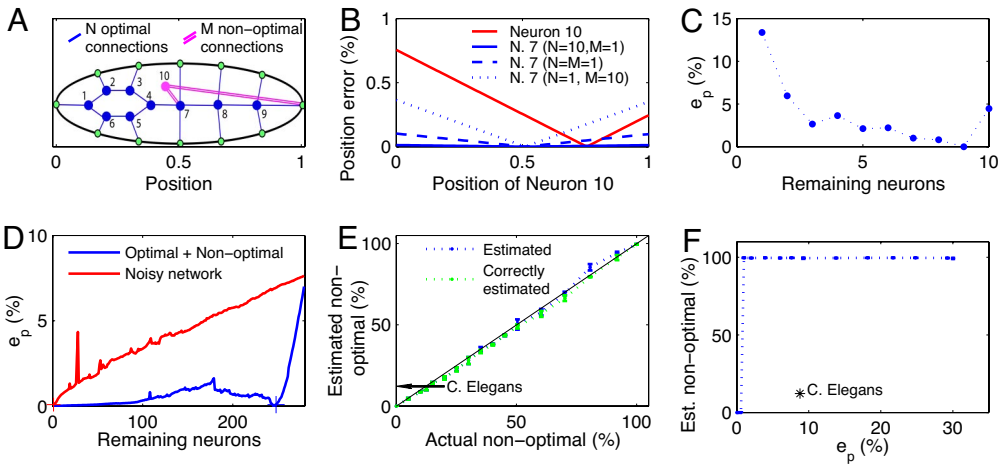


Fig. 4. Method to dissect networks into optimal and nonoptimal subnetworks. (*A*) Toy network configuration. Neurons 1–9 (blue) are optimal, and neuron 10 (pink) is located at random. Blue and pink links represent *N* and *M* connections each, respectively. (*B*) Position error for nonoptimal neuron 10 (red) and optimal neuron 7 (blue) as a function of the position of the nonoptimal neuron 10 (rest of neurons have vanishing error). Step 1 in the dissection method classifies neurons by their position error: neuron 10 is classified as the worst one when the red line is above the blue line ($N > M$ and neuron 10 located sufficiently far from its center-of-mass position), then neuron 7, and then the rest. (*C*) Step 2 in the dissection method calculates the average position error e_p in networks of decreasing size (from right to left) eliminating worst located neurons in the order determined in step 1 of the method. Minimum of the error located at size 9. (*D*) Same as *C* but for a network configuration with the *C. elegans* connectivity and with optimal and nonoptimal subnetworks (blue) and for a noisy network with neurons located at positions obtained adding to optimal positions noise following a uniform distribution of std = 8% (red). Plus signs indicate the point separating optimal and nonoptimal subnetworks. (*E*) Estimated size of the nonoptimal subnetwork (blue) and percentage of correctly estimated nonoptimal neurons (green). (*F*) Dissection method does not find a separation between optimal and nonoptimal components in noisy networks built without optimal and nonoptimal components.

connectivity illustrated in Fig. 4*A*, with 9 optimal neurons in blue and the 10th neuron in pink in a random position, also connected to sensors or muscles (green circles). Blue and pink links indicate N and M connections, respectively. The first step of the method consists in classifying the neurons by their position error measured for each neuron as the difference between its actual position and its locally optimal position, calculated as the position of its center-of-mass that in general reads as

$$x_i^{CM} = \frac{\alpha \sum_{j \neq i} A_{ij} x_j^{\text{actual}} + \sum_k B_{ik} S_k + \beta \sum_l C_{il} m_l}{\alpha \sum_{j \neq i} A_{ij} + \sum_k B_{ik} + \beta \sum_l C_{il}}, \qquad [5]$$

with x_j^{actual} the actual positions of the rest of the neurons ($\alpha = \beta = 1$ in toy model). For the example in Fig. 4*A*, all of the optimally placed neurons have vanishing position errors, except neuron 7 that is connected also to the nonoptimal neuron 10. Although neuron 7 is optimal, its estimated position error depends linearly on the position of neuron 10 with a slope given by $M/(M + 4N)$. As long as the nonoptimal neuron 10 is located sufficiently far from its center-of-mass position and the optimal links outnumber nonoptimal ones, $N\,M$, neuron 7 is then correctly classified as having less position error than neuron 10, (Fig. 4*B*, continuous and dashed blue line). The second step of the method consists in performing a full optimization using Eq. **2** with the complete network first and then further optimization calculations eliminating one by one the neurons in order of decreasing position error as determined by step 1, and for each of these subnetworks we calculate the average position error e_p. In our example this procedure corresponds to an optimization calculation with all neurons first, then we eliminate neuron 10, followed by neuron 7 and then any of the rest, and calculate for each case the average position error, Fig. 4*C*. The lowest average position error is found for neurons 1–9, that is, the method finds the optimal subnetwork. Two lessons can be extracted from this simple example. First, this simple method works better the lower the number of nonoptimal links. Second, it is clear that step 1 can be improved by using an iterative classification. Once neuron 10 is classified as the worst one, the remaining 9 are reclassified. From these 9 neurons the worst neuron is selected and put as second worse in the list. This reclassification process is continued until the last neuron. In this way optimal neurons linked to nonoptimal ones, like neuron 7, are not necessary classified as worse than the rest by the dissection method. In the rest of the paper we use this iterative step 1 unless stated otherwise.

We then applied the dissection method to artificial network configurations with the same connectivity as the *C. elegans* network to test its efficacy in separating optimal and nonoptimal subnetworks. We built network configurations with a given percentage of randomly selected nonoptimal neurons with positions taken from a uniform distribution of width given by the length of the animal. The rest of the neurons are optimal, that is, obtained from an optimization calculation including only these neurons. We applied to these network configurations the iterative step 1 to classify the neurons by position error and step 2 to calculate the average position error e_p in subnetworks of decreasing size by gradually eliminating the worst neurons in the list of step 1 (Fig. 4*D*, blue curve). The average position error e_p goes down to a value very close to zero (as seen from right to left) at a network size that separates the optimal and nonoptimal subnetworks. The dissection method was found to have an excellent performance in the estimation of the size of the nonoptimal network for nonoptimal subnetworks up to a size of 50% and deteriorates only slightly for larger nonoptimal networks (Fig. 4*E*, blue). The percentage of neurons correctly classified as nonoptimal (truly nonoptimal neurons in the estimated nonoptimal subnetwork) is also excellent with little misclassification until a size of 50% of the total network and a slight deterioration for increasing size (Fig. 4*E*, green). Similar results are found for neurons in nonoptimal networks located randomly following a Gaussian instead of a uniform distribution (SI Fig. 7*B*). Crucial for the success of the method is the iterative step 1, without which the deterioration of the estimation already starts at a size of nonoptimal subnetwork of 25% (SI Fig. 8*B*).

Networks formed by a near-optimal subnetwork with a given low noise added to optimal positions and a nonoptimal subnetwork with larger noise are also dissected to a good approximation when the nonoptimal subnetwork has a size <50% of the total network (SI Fig. 9*B*). The distinct feature of the average position error e_p in step 2 for this case is not a minimum but an abrupt change in the slope when reducing the size of the network. This point of abrupt slope change separates well the optimal from the nonoptimal subnetworks (SI Fig. 9*A*). We have implemented a robust numerical algorithm to find this separation point automatically (*SI Text*). Noisy networks, that is, networks in which neurons are randomly displaced from its optimal location but do not have optimal and nonoptimal subnetworks, also have lower average position error e_p when eliminating the worst located neurons. Interestingly, however, these noisy networks never show, whatever the noise level, a special point that could be used to separate optimal from nonoptimal subnetworks like a minimum or a change in slope (Fig. 4*D*, red line for uniform noise and SI Fig. 9*A*, red line for Gaussian noise). As a consequence, our algorithm classifies noisy networks either as fully optimal when the noise is very low or as fully nonoptimal for higher noise (Fig. 4*F* for uniform noise and SI Fig. 9*C* for Gaussian noise).

Dissection of the actual *C. elegans* network results in a nonoptimal subnetwork of 12% of neurons (34 of 279) with 20 interneurons, 8 sensory neurons and 6 motor neurons. The average position error e_p is lowered from 8.8% for the complete network to 5.4% for the near-optimal subnetwork (Fig. 5*A*, see SI Table 1 for a list of the nonoptimal neurons). More importantly, the clustering error e_c is lowered very significantly from 14% to 2.1% (Fig. 5*B*) so the clustering diagram of the optimal subnetwork (Fig. 5*B*, *Large Inset*) is now very similar to the experimental one down to fine details like the clustering in ganglia within the brain (Fig. 1*D*). Encephalization, for example, is explained by using wiring economy for the optimal subnetwork (SI Fig. 10). By using the optimal subnetwork and a

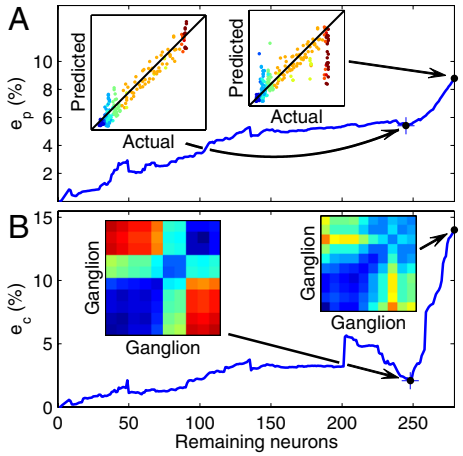


Fig. 5. Dissection of the actual *C. elegans* configuration into near-optimal and nonoptimal subnetworks. (*A*) Average position error e_p in networks of decreasing size (from right to left) eliminating worst located neurons in the order determined in step 1 of the method. Separation of optimal and nonoptimal subnetworks was found automatically at the point with a plus sign corresponding to a large slope change. Insets: Predicted versus actual position for the near-optimal subnetwork and for the complete network. (*B*) Same as *A* but for the clustering error. (*Insets*) Clustering diagram for near-optimal subnetwork (compare with actual one in Fig. 1*D*) and for total network.

simple model for synapse location at the middle point between somas, predicted location of synapses is found to be consistent with the actual nerve ring (SI Fig. 11). Note that the error of the near-optimal subnetwork in *C. elegans* does not go to zero, consistent with a network in which both the near-optimal and nonoptimal subnetworks have noise of different variances (SI Fig. 9*D*).

Wiring economy applied to the near-optimal subnetwork in *C. elegans* accounts for the clustering of neurons into ganglia as shown by the low clustering error of $e_c = 2.1\%$. The origin of this low value is 2-fold. First, the dissection method has eliminated badly located nonoptimal neurons thus reducing average position error. Second, in the optimization of the complete network, the nonoptimal subnetwork is not passive but has a negative effect on the near-optimal subnetwork. The clustering error of the near-optimal subnetwork in the complete network optimization is found to be 3% higher than alone. To further understand this effect, we have searched for the nonoptimal neurons responsible for increased position errors in the near-optimal neurons when both are forming the complete network. We removed one neuron at a time and calculated the change in position error in each of the near-optimal neurons (SI Fig. 12*A*). Neuron labels in the figure are located such that most damaging nonoptimal neurons are at the bottom and most affected near-optimal neurons to the left. The most damaging nonoptimal neurons are the interneurons AVA, DVA, PVC and AKL. The near-optimal neurons most affected by these nonoptimal neurons are PHC and PHB, but many more are affected and this effect is in part responsible for the low clustering error when eliminating the nonoptimal subnetwork. Nonoptimal neurons also affect the position error of other nonoptimal neurons, but the effects are smaller than on near-optimal neurons and in this case there is no clear core of most affecting neurons (SI Fig. 12*B*). Only DVA appears again as one of the more damaging neurons.

Discussion

The *C. elegans* nervous system was found to contain a subnetwork of 84% of neurons wired almost optimally. This subnetwork explains the origin of the structure of the whole system, in particular the segregation of neurons into ganglia and encephalization. A small subset of the connections of this large near-optimal subnetwork formed by the links of sensory and motor neurons to sensors and muscles is responsible for the main features of the neuroanatomy. Connections among neurons are responsible for further neuroanatomical detail, like the formation of ganglia within the head ganglia and the location of this head ganglia within the body. An analysis of nonoptimal neurons revealed that interneurons AVA, DVA, PVC, and AKL have a negative effect on the position of near-optimal neurons in a complete network optimization. This result is consistent with the idea that these interneurons might be spatially constrained in the actual nematode. Interestingly, DVA is a stretch receptor neuron expressing TRP-4, the *elegans* homologue of the mechanosensitive TPRN channel (36). AVA and PVC express UNC-8, a DEG/EnaC family member homologue to subunits of a mechanically gated ion channel (37). Internal body sensors like those implicated in propioception could be added to the theory in a form identical to external sensors when experimental data becomes available.

Materials and Methods

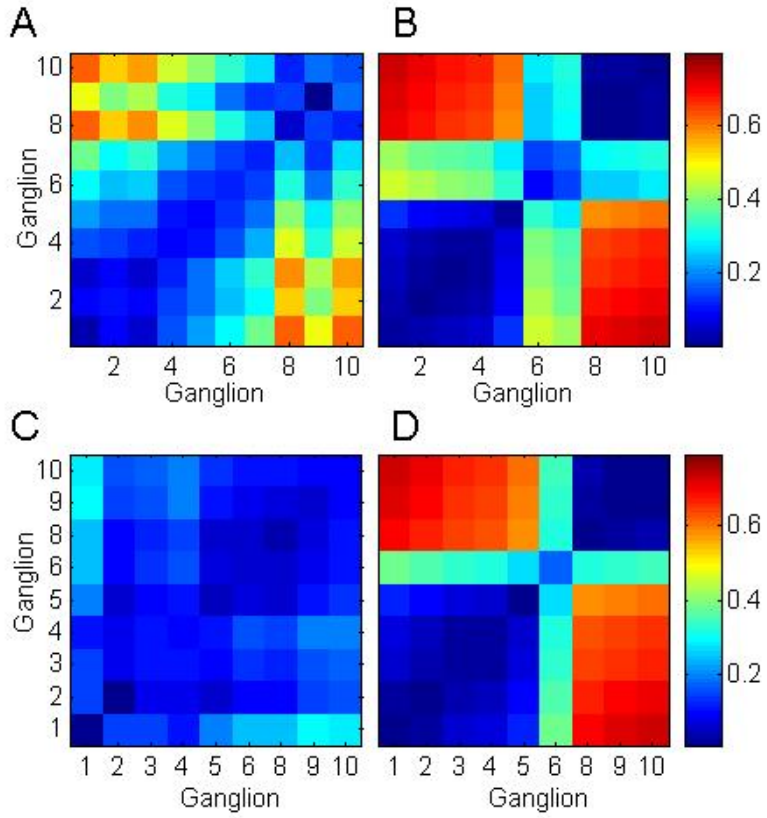
Data. Network connectivity and actual positions of somas, sensors, and muscles were taken from the revised data published in ref. 24, available at www.wormatlas.org.

Construction of Interganglia Distance Graphs. Interganglia distance graphs, or clustering graphs, as those in Fig. 1 *D* and *E*, were built in the following way. Neurons were grouped by the ganglion they belong to in the real nematode. Then, distances between every neuron of ganglion i and every neuron of ganglion j are computed (or, if $i = j$, between every two neurons of the same ganglion). The average of these distances is the i,j-th element of the interganglia distance graph. Note that these matrices are symmetrical by construction.

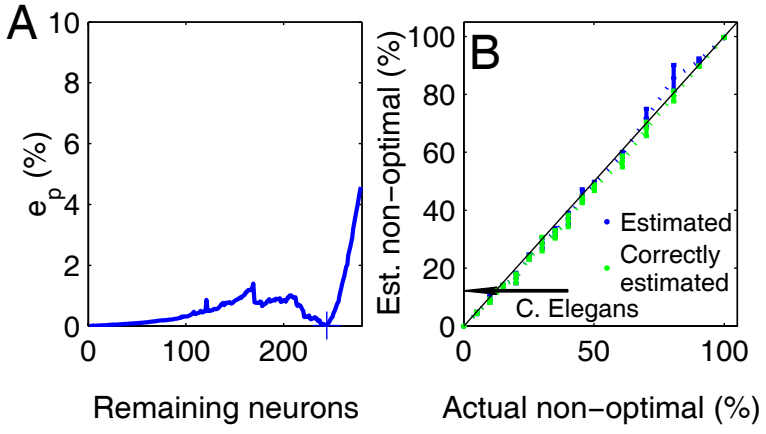
Construction of Interganglia Connectivity Graphs. Square (i,j) of Fig. 1*F* represents the number of connections per neuron between ganglia i and j. It was computed by dividing the total number of connections between ganglia i and j by the sum of the number of neurons of the two ganglia. When $i = j$, it is computed as the total number of connections between neurons of the same ganglion divided by the number of neurons that form the ganglion. For Fig. 1*G*, the length of the nematode was divided into 10 bins of equal length. Square (i,j) of this figure represents the number of connections between ganglion i and the organs that fall into the j-th bin, divided by the number of neurons in ganglion i.

We thank Brian Burton, Vicente Pérez-Díaz, and Adela Escudero-Berián for critical comments on the manuscript. This work was supported by the Spanish Ministry of Science and Education (MEC) (G.G.d.P.) and Comunidad de Madrid (Biociencia program) (G.G.d.P.). A.P.-E. acknowledges fellowships from Consejo Superior de Investigaciones Cientificas, MEC, and Instituto de Salud Carlos III.

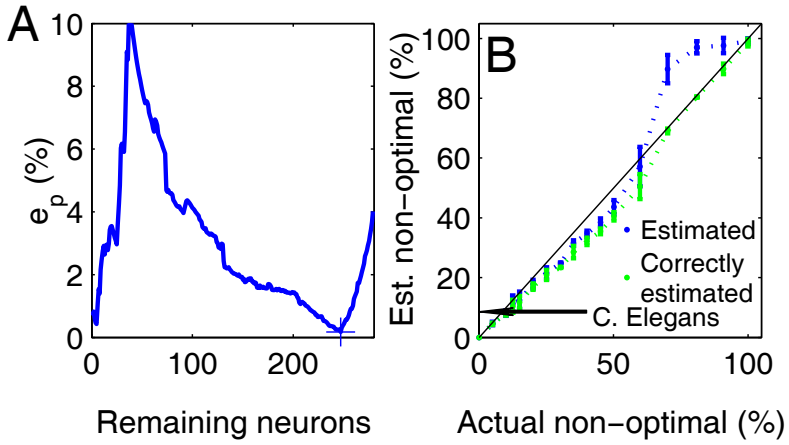
1. Striedter GF (2004) *Principles of Brain Evolution* (Sinauer, Sunderland, MA).
2. Nusbaum MP, Blitz DM, Swensen AM, Wood D, Marder E (2001) *Trends Neurosci* 24:146–154.
3. Konishi M (1999) in *Elsevier's Encyclopedia of Neuroscience*, eds Adelman G, Smith BH (Elsevier, Amsterdam), pp 1906–1908.
4. Ramón y Cajal S (1899) *Textura del Sistema Nervioso del Hombre y de los Vertebrados* (Nicolás Moya, Madrid) (Spanish); trans Pasik P, Pasik T (1999) *Texture of the Nervous System of Man and the Vertebrates* (Springer, New York).
5. Mitchison G (1991) *Proc R Soc London B Biol Sci* 245:151–158.
6. Van Essen DC (1997) *Nature* 385:313–318.
7. Cowey A (1979) *Q J Exp Psychol* 31:1–17.
8. Nelson ME, Bower JM (1990) *Trends Neurosci* 13:403–408.
9. Chklovskii DB, Koulakov A (2000) *Physica A* 284:318–334.
10. Durbin R, Mitchison G (1990) *Nature* 343:644–647.
11. Koulakov AA, Chklovskii DB (2001) *Neuron* 29:519–527.
12. Murre JM, Sturdy DP (1995) *Biol Cybern* 73:529–545.
13. Ruppin E, Schwartz EL, Yeshurun Y (1993) *Biol Cybern* 70:89–94.
14. Wen Q, Chklovskii DB (2005) *PLoS Comput Biol* 1:e78.
15. Cherniak C (1994) *J Neurosci* 18:522–527.
16. Cherniak C (1995) *Trends Neurosci* 14:2418–2427.
17. Klyachko VA, Stevens CF (2003) *Proc Natl Acad Sci USA* 100:7937–7941.
18. Chklovskii DB (2004) *Neural Comput* 16:2067–2078.
19. Cherniak C, Changizi M, Won Kang D (1999) *Phys Rev E* 59:6001–6009.
20. Chklovskii DB (2000) *J Neurophysiol* 83:2113–2119.
21. Chklovskii DB, Stepanyants A (2003) *BMC Neuroscience* 4:18.
22. Cherniak C (1992) *Biol Cybern* 66:503–510.
23. Shefi O, Harel A, Chklovskii DB, Ben-Jacob E, Ayali A (2003) *Neurocomputing* 58–60:487–495.
24. Chen BL, Hall DH, Chklovskii DB (2006) *Proc Natl Acad Sci USA* 103:4723–4728.
25. White JG, Southgate E, Thomson JN, Brenner S (1986) *Philos Trans R Soc London Ser B* 314:1–340.
26. Hall DH, Russell RL (1991) *J Neurosci* 11:1–22.
27. Hall K (1970) *Management Sci* 17:219–229.
28. Kleinhaus J, Sigl G, Johannes F, Antreich K (1991) *IEEE Trans CAD* 10:356–365.
29. Ahn, Y-Y, Jeong H, Kim BJ (2006) *Physics A* 367:531–537.
30. Kaiser M, Hilgetag CC (2006) *PLoS Comput Biol* 2:e95.
31. Rushton WA (1951) *J Physiol (London)* 115:101–122.
32. Rall W, Burke RE, Holmes WR, Jack JJ, Redman SJ, Segev I (1992) *Physiol Rev* 72:5159–5186.
33. Laughlin SB, de Ruyter van Steveninck R, Anderson JC (1998) *Nat Neurosci* 1:36–41.
34. Attwell D, Laughlin SB (2001) *J Cereb Blood Flow Metab* 21:1133–1145.
35. Tessier-Lavigne M, Goodman CS (1996) *Science* 274:1123–1133.
36. Li W, Feng Z, Sternberg PW, Xu XZ (2006) *Nature* 440:684–687.
37. Tavernarakis N, Shreffler W, Wang S, Driscoll M (1997) *Neuron* 18:107–119.



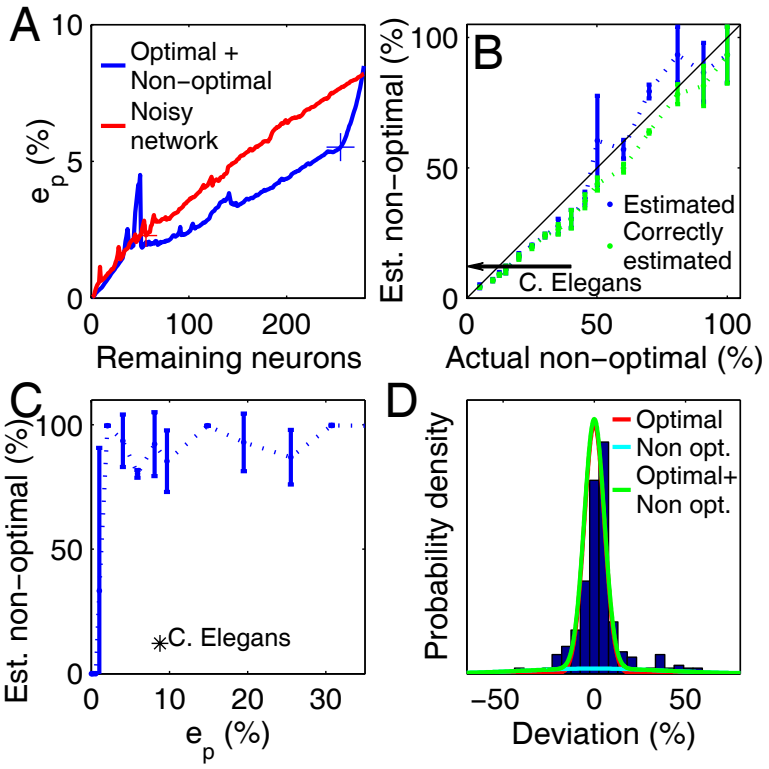
SI Figure 6: Clustering of sensory and motorneurons and interneurons in an optimization of full network is different to actual clustering. (*A*) Average distance between neurons belonging to the same ganglion (diagonal elements) or different ganglia (nondiagonal elements) for sensory and motorneuron positions when the complete network is optimised and $a = b = 1/29.3$. Clustering error is $e_c = 12.3\%$. (*B*) Same as A but for the actual nematode. (*C*) Same as A but for interneurons. Ganglion 7 has been omitted because it contains no interneurons. Clustering error is $e_c = 25.2\%$. (*D*) Same as *C* but for the actual nematode.



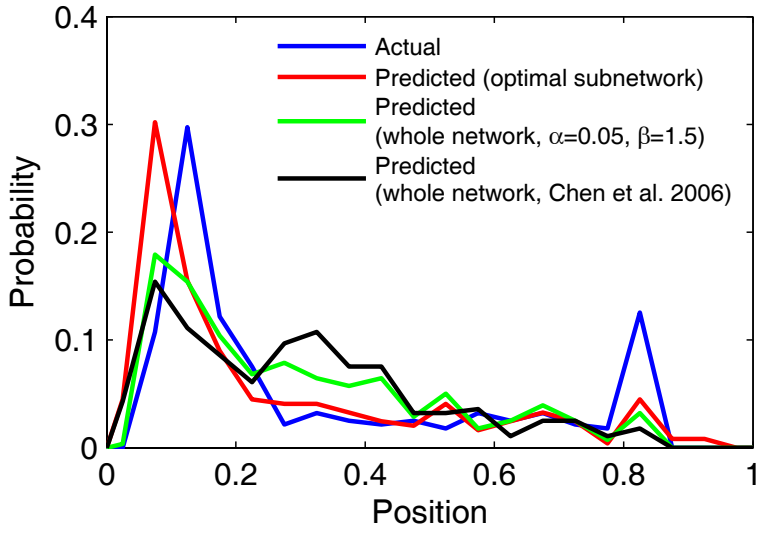
SI Figure 7: Method to dissect networks with noiseless optimal subnetwork and nonoptimal subnetworks with Gaussian noise. (*A*) Step 2 in the dissection method calculates the average position error in networks of decreasing size eliminating worst located neurons in the order determined in step 1 of the method. Position of somas for optimal subnetwork obtained from an optimization of that subnetwork. Soma positions for the nonoptimal subnetwork obtained from an optimization of the complete network and then adding Gaussian noise of std = 35% to the neurons selected as nonoptimal. Point with plus sign indicates the point that our algorithm finds separating optimal and nonoptimal subnetworks. (*B*) Estimated size of the nonoptimal subnetwork (blue) and correctly estimated nonoptimal neurons (green).



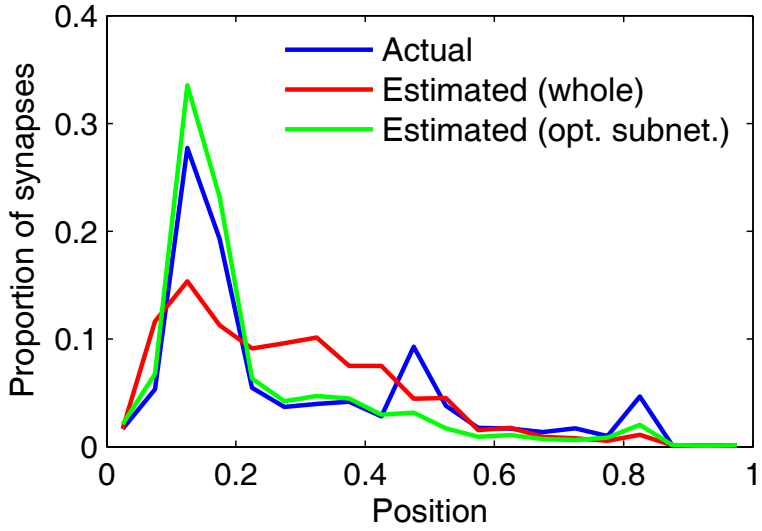
SI Figure 8: Same as SI Fig. 7 but Step 1 in the dissection method does not classify neurons using an iterative method



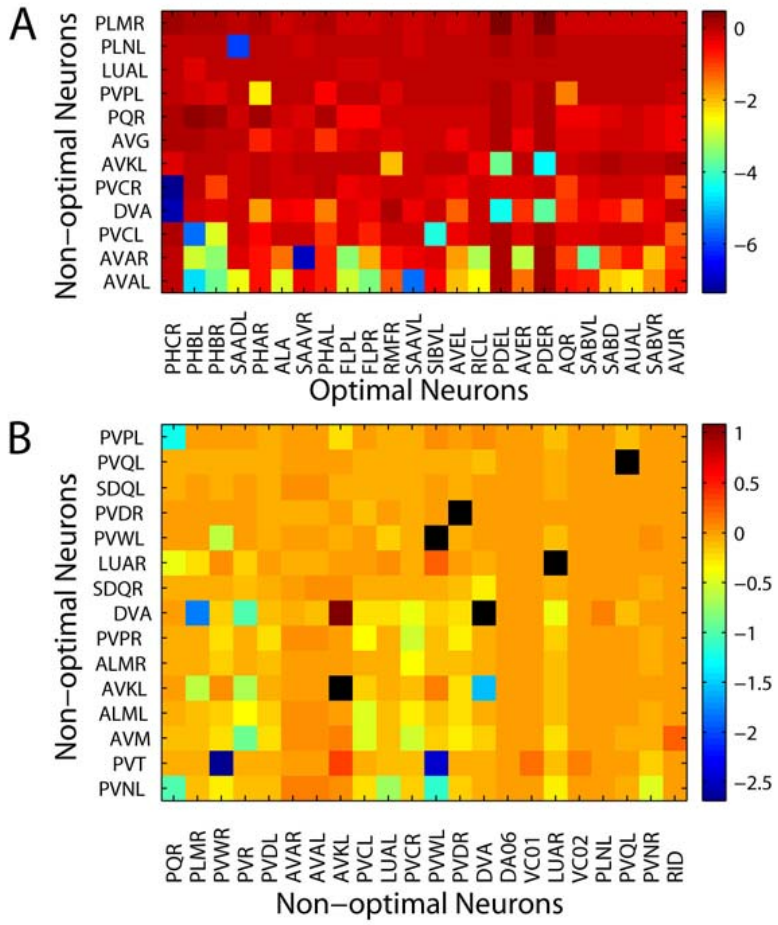
SI Figure 9: Method to dissect networks with near-optimal subnetwork with low Gaussian noise (std = 5.5%) and nonoptimal subnetworks with larger Gaussian noise (std = 35%). (*A*) Step 2 in the dissection method calculates the average position error in networks of decreasing size eliminating worst located neurons in the order determined in step 1 of the method. Blue line, case of network with near-optimal and nonoptimal subnetworks. Position of somas for near-optimal subnetwork obtained from an optimization of that subnetwork and adding Gaussian noise of std = 5.5% (as obtained in the elegans near-optimal subnetwork). Soma positions for the nonoptimal subnetwork obtained from an optimization of the complete network and then adding Gaussian noise of std = 35% to the neurons selected as nonoptimal. Red line, noisy network with neuron positions obtained adding Gaussian noise of std = 8% to the optimal position of all neurons. (*B*) Estimated size of the nonoptimal subnetwork (blue points) and correctly estimated nonoptimal neurons (green points) for networks formed by a near-optimal subnetwork and nonoptimal subnetworks of different sizes. (*C*) Same as *B* but for noisy networks of different std. (*D*) Deviation of neuron position from optimal position in actual nematode consistent with near-optimal subnetwork of Gaussian noise of std = 5.5% and nonoptimal subnetwork of std = 35%.



SI Figure 10: Encephalization in actual network and optimal subnetwork and optimization of complete network. Histogram of number of neurons along the animal.



SI Figure 11: Histogram of synapse location in the animal assuming that synapses are at midpoints between soma.



SI Figure 12: Interaction matrices for nonoptimal and near-optimal neurons. (*A*) Change in position error in each of the near-optimal neurons (bottom) when removing one of the nonoptimal neurons (left) from the network. Neuron labels in the figure are positioned such that most damaging nonoptimal neurons are at the bottom and most damaged near-optimal neurons to the left. Only the 12×24 matrix at the bottom-left of the complete 34×184 matrix (SI Fig. 13) is shown here. (*B*) Same as *A* but for the change in position error of nonoptimal neurons. Shown are the 15×22 of most damaging and most damaged neurons (see SI Fig. 14 for the complete 34×34 matrix).

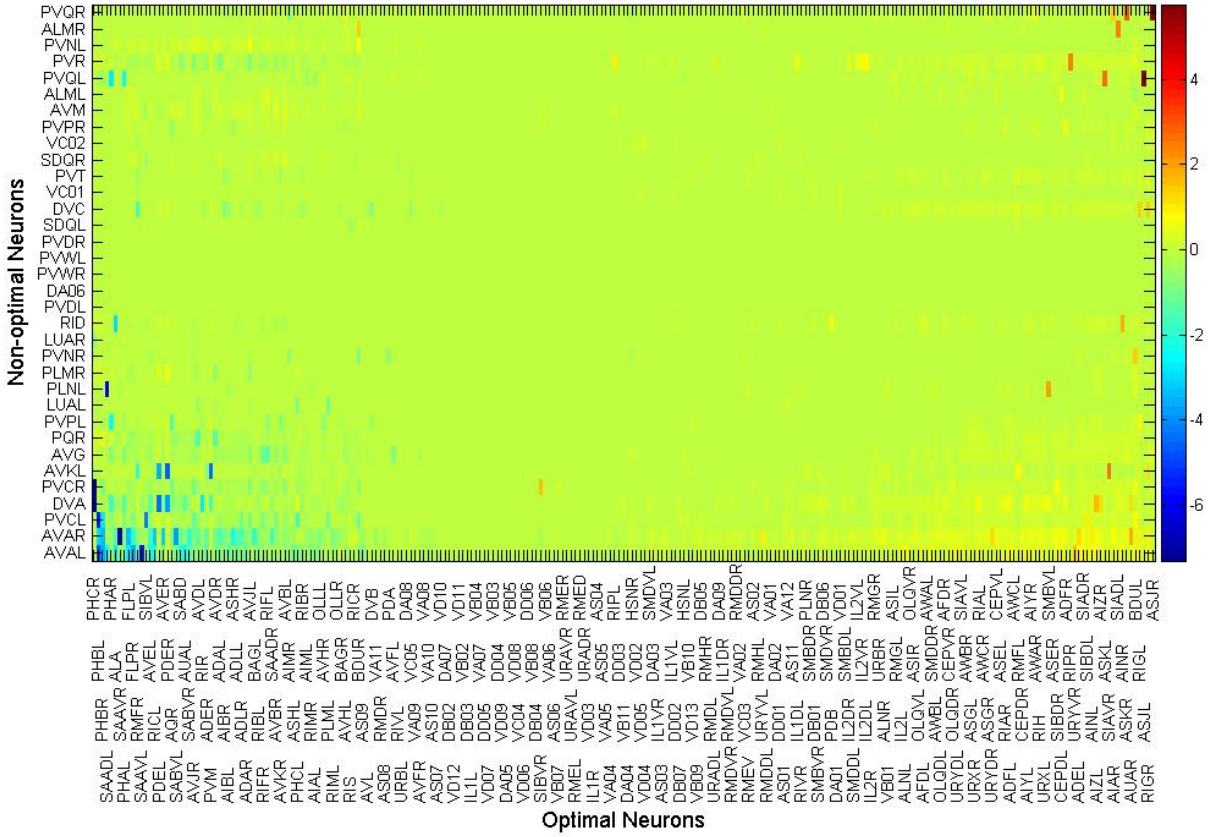


Figure 13: Interaction matrix for nonoptimal and near-optimal neurons. Shown is the change in position error in the near-optimal neurons (bottom) when removing one of the nonoptimal neurons (left) from the complete network. Neuron labels in the figure are positioned such that most damaging nonoptimal neurons are at the bottom and most affected near-optimal neurons are to the left.

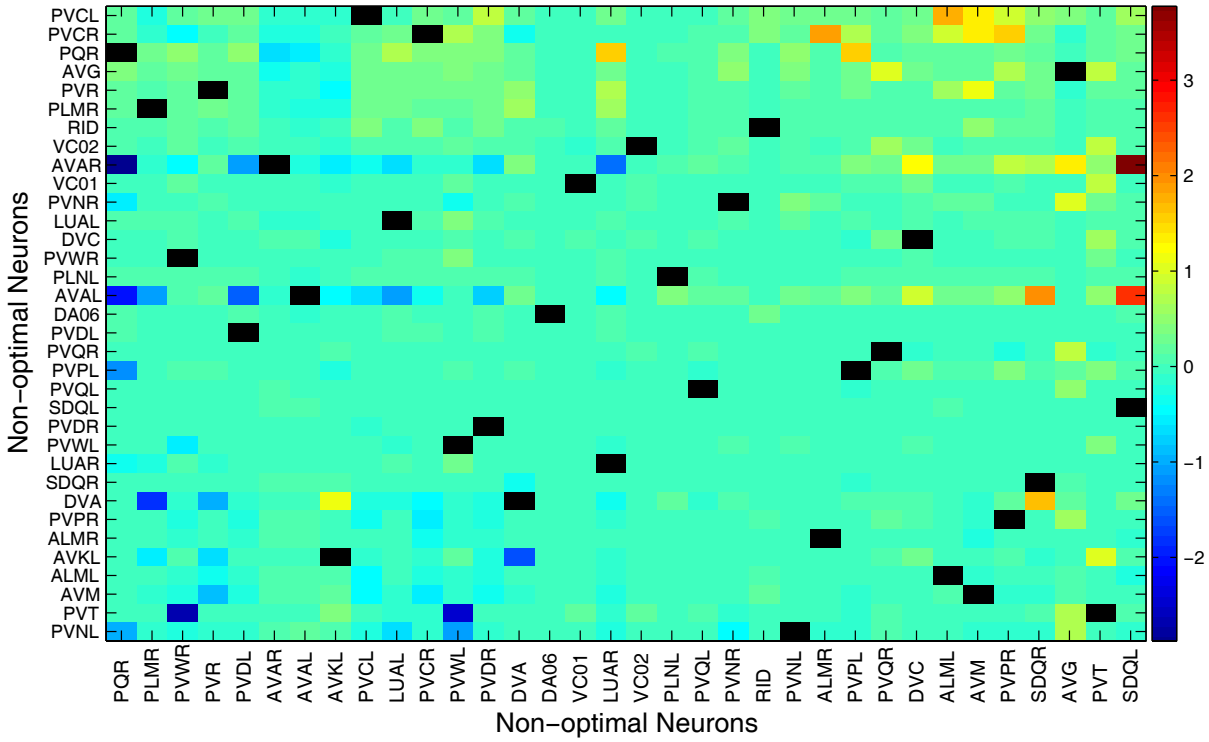


Figure 14: Interaction matrix between nonoptimal neurons. Shown is the change in position error in the nonoptimal neurons (bottom) when removing one of the nonoptimal neurons (left). Neuron labels in the figure are positioned such that most damaging nonoptimal neurons are at the bottom and most affected nonoptimal neurons are to the left.

Index	Name	Type
160	PVQL	Interneuron
161	PVQR	Interneuron
156	PVNL	Motor
128	LUAL	Interneuron
163	PVT	Interneuron
172	RID	Motor
157	PVNR	Motor
158	PVPL	Interneuron
111	DVC	Interneuron
159	PVPR	Interneuron
64	AVG	Sensory
209	SDQL	Interneuron
129	LUAR	Interneuron
152	PVCR	Interneuron
151	PVCL	Interneuron
153	PVDL	Interneuron
154	PVDR	Interneuron
150	PQR	Sensory
165	PVWR	Interneuron
164	PVWL	Interneuron
109	DVA	Interneuron
162	PVR	Sensory
147	PLMR	Sensory
54	AVAL	Interneuron
55	AVAR	Interneuron
92	DA06	Motor
24	ALML	Sensory
72	AVM	Sensory
25	ALMR	Sensory
69	AVKL	Interneuron
210	SDQR	Interneuron
262	VC01	Motor
148	PLNL	Sensory
263	VC02	Motor

SI Table 1: List of the most nonoptimal neurons in order of decreasing nonoptimality

SI Text

This Supporting Information has been copied verbatim from the version published in *PNAS*, except for some typographical errors that have been corrected.

References

1. Hall K (1970) Management Sci 17:219-229
2. Chklovskii DB (2004) Neural Comput 16:2067-2078

Derivation of Formula for Positions Minimizing Wiring Cost, Eq. 2. In the following we give for completeness a very explicit derivation of Eq. 2. Equivalent derivations based on a heavier use of matrix properties can be found in references 1 and 2. When cost increases quadratically with wire length, total cost is given by Eq. 1 with $x = 2$,

$$W = \frac{1}{2}\alpha \sum_{i,j=1}^{N} A_{ij} \left(x_i - x_j\right)^2 + \sum_{i,k=1}^{N,S} B_{ik} \left(x_i - s_k\right)^2 + \beta \sum_{i,l=1}^{N,M} C_{ik} \left(x_i - m_l\right)^2$$

This function has a minimum where partial derivatives with respect to the positions of all neurons are zero,

$$\frac{\partial W}{\partial x_p} = 0,$$

for all $p = 1, \ldots, N$, with

$$\begin{aligned}\frac{\partial W}{\partial x_p} =& \frac{\alpha}{2} \sum_{i,j=1}^{N} A_{ij} \left[2\delta_{ip}(x_i - x_j) - 2\delta_{jp}(x_i - x_j)\right] \\ &+ \sum_{i,k=1}^{N,S} B_{ik} 2\delta_{ip}(x_i - s_k) + \beta \sum_{i,l=1}^{N,M} C_{il} 2\delta_{ip}(x_i - m_l),\end{aligned}$$

and δ_{ij} the Kronecker delta ($\delta_{ij} = 1$ when $i = j$ and 0 otherwise). Because of the Kronecker delta functions, most terms vanish. However, we keep most of them for later convenience. After regrouping terms and removing some of the vanishing terms, we get

$$\begin{aligned}\alpha \sum_{i,j=1}^{N} A_{ij}\delta_{ip}(x_p - x_j) + \alpha \sum_{i,j=1}^{N} A_{ij}\delta_{jp}(x_j - x_i) + 2 \sum_{i,k=1}^{N,S} B_{ik}\delta_{ip}x_i \\ + 2\beta \sum_{i,l=1}^{N,M} C_{il}\delta_{ip}x_i = 2 \sum_{k=1}^{S} B_{pk}s_k + 2\beta \sum_{l=1}^{M} C_{pl}m_l.\end{aligned}$$

As the matrix A is symmetric, the first and second terms are identical, giving

$$\begin{aligned}\alpha \sum_{i,j=1}^{N} A_{ij}\delta_{ip}(x_p - x_j) + \sum_{i,k=1}^{N,S} B_{ik}\delta_{ip}x_i + \beta \sum_{i,l=1}^{N,M} C_{il}\delta_{ip}x_i = \\ \sum_{k=1}^{S} B_{pk}s_k + \beta \sum_{l=1}^{M} C_{pl}m_l.\end{aligned}$$

Regrouping terms we obtain[1]

$$\begin{aligned}\sum_{i=1}^{N} \delta_{ip} \left(\alpha \sum_{j=1}^{N} A_{pj} + \sum_{k=1}^{S} B_{ik} + \beta \sum_{l=1}^{M} C_{pl}\right) x_p - \alpha \sum_{j=1}^{N} A_{pj}x_i = \\ \sum_{k=1}^{S} B_{ik}s_k + \beta \sum_{l=1}^{M} C_{pl}m_l,\end{aligned}$$

where we have eliminated vanishing terms in the second term. Renaming the summation index j for i in the second term, and using

[1] A typographical error in this equation has been corrected with respect to the version published in *PNAS*

the definition of Q in Eq. 2b, we can write

$$\sum_{i=1}^{N} Q_{ip} x_p = \sum_{k=1}^{S} B_{pk} s_k + \beta \sum_{l=1}^{M} C_{pl} m_l; \qquad p = 1 \dots N$$

This is a system of N equations with the soma positions x_i as the N unknowns. In matrix notation we can write it as

$$Q\vec{x} = B\vec{s} + \beta C\vec{m},$$

where $\vec{x}$, $\vec{s}$ and $\vec{m}$ are column vectors that store the positions of all neurons, sensors and muscles, respectively. Multiplying from the left by Q^{-1} both members of the matrix equation, we obtain the solution

$$\vec{x} = Q^{-1} \left[B\vec{s} + \beta C\vec{m} \right],$$

as given in Eq. 2a.

Derivation of the Center of Mass Formula, Eq. 4. When only connections between neurons and organs (sensors and muscles) are considered, and quadratic cost per unit length is assumed, total cost in Eq. 1 reduces to

$$W = \sum_{i,k=1}^{\hat{N},S} B_{ik} (x_i - s_k)^2 + \beta \sum_{i,l=1}^{\hat{N},M} C_{il} (x_i - m_l)^2,$$

where $\hat{N}$ is the number of sensory and motor neurons. The minimum is characterized by vanishing partial derivatives

$$\frac{\partial W}{\partial x_p} = 2 \sum_{k=1}^{S} B_{pk} (x_p - s_k) + 2\beta \sum_{l=1}^{M} C_{pl} (x_p - m_l) = 0.$$

Note that the neuron positions x_p can be taken out from the sums and Eq. 4 in the main text is then obtained directly.

Robustness of Predictions Using the Center-of-Mass Formula, Eq. 4. First, the prediction is almost independent of any parameters. Although the parameter β appears in Eq. 4, only 13 out of 199 neurons are both sensory and motorneurons, and for the remaining 186 neurons the equation further reduces to separate expressions for sensory and motorneurons, which are independent of β. In practice, predictions improve for increasing β[2] but, in any case, the maximum difference is $\Delta e_p = 0.5\%$ for the mean error position and $\Delta e_c = 0.2\%$ for the clustering error. For different powers $\xi = 1.1, \dots, 10$ for the wire cost in Eq. 3, we found differences in position and clustering errors to be $\Delta e_p = 0.4\%$ and $\Delta e_c = 0.03\%$, respectively (see next section below). Even when changing the connectivity matrix B to a matrix of 1s and 0s, the difference was found to be $\Delta e_p = 0.1\%$ for the mean error position and $\Delta e_c = 0.01\%$ for clustering error. The reason for this strong robustness is that, although some of the patches of sensors and muscles to which each ganglion is connected are relatively large, each neuron connects to a much smaller mini-patch of sensors or muscles.

[2] This symbol has been corrected with respect to the version published in *PNAS*

Fig. 2F in the main text gives the statistics of the size of the mini-patch a single neuron connects to, with a maximum at zero length (single connections) and another maximum at a length of 5% the total length of the animal. Any sensible wiring economy approach for this subnetwork would predict the location of each neuron somewhere within the length of the mini-patch it connects to. Therefore, any modification of the model's parameters should cause position differences below 5% the length of the animal. As mentioned above, on average, these differences are below 1%, and affect the clustering error less than 0.3% (adding up all the contributions listed above, due to changing parameters β and ξ and the detailed form of matrix B).

Numerical Wire Cost Minimization for Nonquadratic Cost. In general, the calculations performed in this article have been for cost increasing quadratically with wire length. However, it appears that the results obtained with quadratic cost would hold even if nonquadratic costs are used. For example, in order to test the robustness of predictions of sensory and motor neurons' positions using the center of mass calculation in Eq. 4, we computed the optimal layouts for costs with exponents ξ from 1.1 to 10 in intervals of 0.1. Note that the total cost of connections between neurons and organs in Eq. 3 can be written as

$$W = \sum_{i=1}^{\hat{N}} \left(\sum_{k=1}^{S} B_{ik} \left| x_i - s_k \right|^{\xi} + \beta \sum_{l=1}^{M} C_{il} \left| x_i - m_l \right|^{\xi} \right) = \sum_{i=1}^{\hat{N}} W_i$$

with W_i the cost associated to neuron i. Therefore, the optimization problem reduces to numerically obtain the optimal position of each neuron separately by minimizing its individual cost. A Newton's minimization algorithm was used for these calculations. This algorithm starts at position 0.5, estimates the position of the minimum from the two first derivatives at that point, and iterates until convergence. As cost functions are convex for exponents greater than 1, convergence to the global minimum is guaranteed.

For the general optimization problem, we tested polynomial costs up to degree 4, restricted to be monotonically increasing and convex. Using a multidimensional Newton's algorithm and a multidimensional greatest-gradient algorithm, we obtained no relevant differences compared with quadratic cost. Therefore, predictions seem to be robust independently of the cost function as long as it is monotonically increasing and convex, and the quadratic case seems to be a good representative function.

Detection of Point Separating Optimal From Nonoptimal Subnetworks in e_p and e_c Plots. In Fig. 5A and SI Fig. 9A the point that one identifies visually as the separation point between the optimal and nonoptimal networks is characterized by separating two regimes of the e_p, one with a very steep slope (at the right of the separation point) and one with a flatter slope (at the left). Algorithms based on the first and sec-

ond derivatives were found to be too sensitive to small fluctuations of the function even after smoothing. An alternative and more robust algorithm is detailed in the following. Start at the last point, and estimate the derivative there. This estimation is done by averaging the numerical derivatives at the 6, 5, 4... 1 final points, and keeping the highest average value. In this way, the estimate is reliable both for small and large nonoptimal subnetworks. Then, "roll down" until a derivative n times smaller is found. We found that $n = 5$ always had a good performance and was used in all calculations presented here. This algorithm can, however, get stacked because of small bumps of the function, so we developed a new one based on the same idea but that is robust to small fluctuations, and its results match almost always with visual inspection. The algorithm is as follows:

Estimate derivative at the end, using the last 6 points.
Set Point of Separation = Number of Neurons
If derivative at the end > 0
Do
For c = 1 to Point of Separation
Slopes(c) = (e_p(Point of Sep.) - e_p(c))/(Point of Sep. - c)
End for
If Max(Slopes) > Derivative at the end / n
Point of Separation=Index(Max(Slopes))
End if
Loop while Max(Slopes) > Derivative at the end / n
End if

3
Deviations from optimality

Que a los libros, como a los hombres, los respetamos y admiramos por sus buenas cualidades, pero sólo los amamos por algunos de sus defectos.

(Just as with men, we admire and respect books for their good qualities; but we can only love them for certain faults that they display)

Santiago Ramón y Cajal
Reglas y consejos
de investigación científica

Structure of deviations from optimality in biological systems

Alfonso Pérez-Escudero[a,b,c,1], Marta Rivera-Alba[a,b], and Gonzalo G. de Polavieja[a,b,1]

[a]Instituto Cajal, Consejo Superior de Investigaciones Científicas, Avenida Dr. Arce 37, 28002 Madrid, Spain; [b]Department of Theoretical Physics and Instituto "Nicolás Cabrera" de Física de Materiales, Universidad Autónoma de Madrid, 28049 Madrid, Spain; and [c]Instituto de Óptica "Daza de Valdés", Consejo Superior de Investigaciones Científicas, Serrano 121, 28006 Madrid, Spain

Edited by Stanislas Leibler, The Rockefeller University, New York, NY, and accepted by the Editorial Board September 20, 2009 (received for review May 18, 2009)

Optimization theory has been used to analyze evolutionary adaptation. This theory has explained many features of biological systems, from the genetic code to animal behavior. However, these systems show important deviations from optimality. Typically, these deviations are large in some particular components of the system, whereas others seem to be almost optimal. Deviations from optimality may be due to many factors in evolution, including stochastic effects and finite time, that may not allow the system to reach the ideal optimum. However, we still expect the system to have a higher probability of reaching a state with a higher value of the proposed indirect measure of fitness. In systems of many components, this implies that the largest deviations are expected in those components with less impact on the indirect measure of fitness. Here, we show that this simple probabilistic rule explains deviations from optimality in two very different biological systems. In *Caenorhabditis elegans*, this rule successfully explains the experimental deviations of the position of neurons from the configuration of minimal wiring cost. In *Escherichia coli*, the probabilistic rule correctly obtains the structure of the experimental deviations of metabolic fluxes from the configuration that maximizes biomass production. This approach is proposed to explain or predict more data than optimization theory while using no extra parameters. Thus, it can also be used to find and refine hypotheses about which constraints have shaped biological structures in evolution.

Caenorhabditis elegans | *Escherichia coli* | evolution | neuroanatomy | optimization

Optimization theory has been widely used to analyze evolutionary adaptation (1–29). Despite their success at explaining important features of many biological systems, optimization principles have been criticized for being an excessive simplification of evolution (30, 31). Major factors in evolution not taken into account by optimization theory are, for example, stochasticity, genetic drift, insufficient time to reach the optimum, the existence of local maxima or insufficient genetic variability (3, 25–31). Practitioners of optimization theory answer to this objection that it is not claimed that biological systems are optimal, but that the optimal configuration is a useful reference to study adaptation in biological systems (3). However, in practical applications the following problem arises: When the real system deviates from the optimum, instead of acknowledging that the system is suboptimal, typically a new optimization principle using more parameters can be given to better fit the data. Although this approach might be justified on the grounds that our objective functions need improvement, the problem is that there is no procedure to distinguish deviations that can be explained by nonadaptive factors like stochasticity or finite time in evolution from those that must be explained by improvement of the objective function.

Here, we test a simple rule for the structure of suboptimal biological systems. This rule implies that the components of the system with lesser impact on the objective function are expected to have a higher probability of deviating from the optimum. We test this theoretical result in the neuroanatomy of the nematode *Caenorhabditis elegans*. We correctly obtain the structure of deviations from the minimum wiring configuration. We develop a significance test to check that experimental deviations from optimality correspond to the theoretical pattern. Also, a Bayesian approach is given to estimate better objective functions using the data, taking into account the deviations from optimality. Finally, we test the theoretical result in the metabolic network of *Escherichia coli*. We find that although maximization of ATP and biomass production gives similar predictions in deterministic optimization, it is biomass production that best explains the deviations from the optimum.

Results

Suboptimal Structure in Biological Systems. Deterministic optimization theory finds the system state $\vec{x}$ with the highest value of the objective function $Z(\vec{x})$. Many factors in evolution, including stochastic effects and finite time, may not allow the system to reach the ideal optimum. In these cases, and given that we typically do not have enough details about these factors, it is still reasonable to expect that the system has a higher probability of reaching a state of high Z. In mathematical terms, we can then write the probability $P(\vec{x})$ of finding the system state $\vec{x}$ as a function f that increases with the objective $Z(\vec{x})$,

$$P(\vec{x}) = f(Z(\vec{x})). \quad [1]$$

We have applied Eq. **1** to systems of many components, such as neurons in a neuronal circuit or chemical fluxes in a metabolic network. In these systems, we typically have that the objective changes differently in the directions of some components than in the direction of others. This is illustrated in Fig. 1*A* for the case of a toy system of two components. The objective function falls more slowly in the direction of x_1, and Eq. **1** implies then a higher probability of deviating from the optimum for component 1. Analogously, in systems of many components, we expect larger deviations in the components with smaller impact on the objective.

Irrespective of the form of function f, Eq. **1** implies a particular structure of the deviations from optimality. This can be seen in Fig. 1*B* for the 2D case, where we illustrate that the isoprobability lines must be identical to the isoobjective lines, independently of the function f. We can thus systematically use Eq. **1** to explain or predict structure of deviations from optimality, with

Author contributions: A.P.-E. and G.G.d.P. designed research; A.P.-E. performed research; A.P.-E., M.R.-A., and G.G.d.P. analyzed data; and A.P.-E. and G.G.d.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. S.L. is a guest editor invited by the Editorial Board.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: gonzalo.polavieja@cajal.csic.es or alfonso.perez.escudero@cajal.csic.es.

This article contains supporting information online at www.pnas.org/cgi/content/full/0905336106/DCSupplemental.

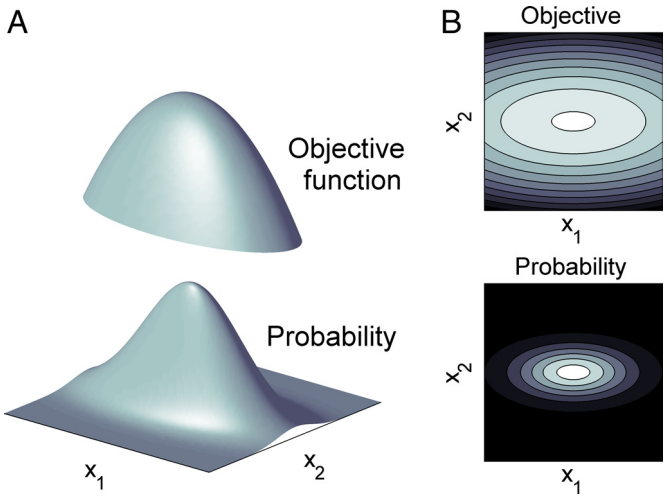


Fig. 1. Suboptimal structure in a two-component model with objective $Z(x_1,x_2) = x_1^2 + 5x_2^2$. (*A*) (*Upper*) Parabolic objective function with slower decrease in the x_1 direction. (*Lower*) Probability resulting from Eq. **1**, using as an example $P(\vec{x}) = \exp(-Z(\vec{x})/0.4)$. The region of high probability extends further in the direction of x_1 because the objective decreases slower in this direction. (*B*) (*Upper*) Contour plots for the objective function (grayscale, lighter for higher values). (*Lower*) Contour plot for the probability (grayscale, lighter for higher values). Eq. **1** implies that isoprobability lines are the same than isoobjective lines.

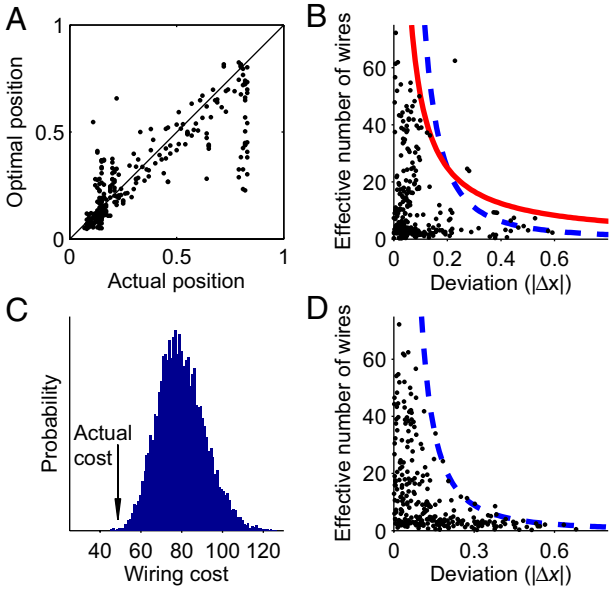


Fig. 2. Deviations of soma positions in *C. elegans* from the optimal positions of minimum wiring configuration. (*A*) Position of somas obtained by deterministic wiring cost minimization versus experimental values. Perfect match between deterministic optimization theory and experiment would fall on the diagonal. (*B*) Effective number of wires of each neuron (ω) versus experimental deviations from optimality ($|\Delta x| \equiv |x_{\text{experimental}} - x_{\text{opt}}|$). Larger deviations are expected for neurons with lower ω. Blue dashed line follows $\omega \propto |\Delta x|^{-2}$, and red solid line follows $\omega \propto 1/|\Delta x|^{-1}$. The three outliers (above the red line) are neurons DA6, AVAL, and AVAR. (*C*) Histogram of the wiring costs resulting from random redistribution of the deviations of somatic positions from their optima. Arrow indicates the cost of the actual configuration. Only 0.033% of the permutations have a lower cost. (*D*) Effective number of wires (ω) versus deviations obtained from a simulation performed by stochastic hill climbing with a Gaussian stochastic component added to wiring cost. Blue dashed line corresponds to the approximate theoretical prediction, $\omega \propto |\Delta x|^{-2}$.

no need of determining a specific function f, and therefore using no more parameters than those of the objective function.

Before using Eq. **1** to explain experimental deviations from optimality, we checked that it corresponds to the resulting probability in simple models of optimization for finite time or in presence of stochastic factors. To model stochasticity, representing either our ignorance of the many unknown extra factors affecting fitness or true system fluctuations, we may consider a stochastic version of the objective, that is the sum of the deterministic objective and a random variable independent of the system variables. The probability of finding an optimum is higher in the direction in which the objective is also higher because in this direction the stochastic component has a higher probability of producing a new global maximum, resulting in Eq. **1** (see *SI Appendix*). For the more complex case with stochasticity present and with the system state changing according to local information of the objective surface, like in stochastic hill climbing (32) (*Materials and Methods*), the probability at equilibrium is again only a function of the objective function, *SI Appendix*. The effect of finite time may be modeled in a simple way as random jumps among system states for a short time and selecting among them the state with highest value of the objective. Again, the probability can only depend on the objective function, Eq. **1** (see *SI Appendix*).

Deviations from Optimality in the Neuroanatomy of *C. elegans*. We then tested whether the organization of *C. elegans* neuroanatomy corresponds to the suboptimal structure given by Eq. **1**. *C. elegans* is a nematode whose nervous system consists of about 300 neurons. The principle of wiring economy states that the neurons will be in the positions that minimize the cost of the wire needed to connect them (refs. 18–21; *Materials and Methods*). Previous theoretical analysis obtained the positions of cell bodies (somas) corresponding to minimum wiring cost, given the observed connectivity (18, 21). The result of this calculation is reproduced in Fig. 2*A*, showing the good agreement between the optimal and the actual positions of the neurons. However, the system is not optimal (18–21), with ≈15% of neurons showing important deviations from optimality. Here, we used the same wiring cost function as in optimization studies to explain the suboptimal structure of the nervous system. From Eq. **1**, deviations from optimality are expected to be larger for the directions of lower cost $W(\vec{x})$ (or equivalently higher objective function, $Z(\vec{x}) = -W(\vec{x})$).

The wiring cost in the direction of each neuron position is found to increase faster the more wires that neuron has

$$W(x_i) - W(x_i^{\text{opt}}) = \omega_i (x_i - x_i^{\text{opt}})^2, \qquad [2]$$

with ω_i a measure of the number of wires associated with neuron i (see *SI Appendix* for a proof of this result). Thus, cost grows parabolically with the distance from the soma to its optimal position x_i^{opt}. Neurons with a lower value of the number of wires ω_i have a slower increasing parabola and are therefore expected to deviate more. This is clearly seen in the experimental data, Fig. 2*B*. All neurons with large deviations from the optimum have a low number of wires. The frontier of the experimental pattern is well described by $\omega \propto 1/|\Delta x|$ for all 279 neurons except the three known as DA6, AVAL, and AVAR (Fig. 2*B*, solid line).

Next, we calculated how significantly these experimental deviations follow the theoretical pattern in Eq. **1** by using the following procedure. Eq. **1** predicts largest deviations to be in the components of the system with lower impact on the cost (neurons with fewer connections). A random reassignation of deviations among neurons is then expected to destroy this effect, increasing the wiring cost. Thus, we randomly redistributed the experimental deviations among the 279 neurons, and calculated how often the cost of the new configuration was lower than the

actual one (*Materials and Methods*). This proportion of cases can be directly treated as a P value, because it is equal to the probability that random deviations get distributed by chance in a way as consistent with our probabilistic model as the empirical data. We found that the experimental deviations correspond very significantly to the probabilistic result, $P = 0.00033$, Fig. 2*C*.

We have also tested that alternative hypotheses cannot explain the experimental data shown in Fig. 2*B*. We checked that the pattern is not due to sampling effects. We also discarded the possibility that experimental errors in the measurement of the positions of neurons, sensors, or muscles are responsible for the structure of the deviations. Finally, we checked that errors in the connectivity matrix cannot explain the data (*SI Appendix*).

Further insight into the experimental deviations in Fig. 2*B* can be obtained from a simulation using the algorithm of stochastic hill climbing with Gaussian noise, Fig. 2*D*. Again, only neurons with low number of wires show important deviations. An approximate theoretical expression for the shape of this pattern can also be obtained from Eq. **1**. Assuming independent neurons, that is, each neuron contributing with a term like Eq. **2** to the total wiring cost, and making the change of variable $u = \sqrt{\omega}\Delta x$, each neuron has the same impact in the wiring cost expressed in the scaled variables and, by Eq. **1**, also in the probability or the cumulative probability. Their isolines then correspond in the old variables to $\omega \propto 1/|\Delta x|^2$ (see *SI Appendix* for a complete proof and tests). As the envelope of the numerical results corresponds to an isoline of the cumulative probability, it follows this expected relationship (Fig. 2*D*, dashed line). The wiring cost function used until now assumes for simplicity that the cost increases with the square of wire length (18, 21) (see *Materials and Methods*). For a wiring cost increasing as wire length to any power ξ, the pattern of deviations approximately obeys $\omega \propto 1/|\Delta x|^{\xi}$, always with largest deviations in neurons with lesser number of wires (*SI Appendix*). For the experimental data in Fig. 2*B*, the quadratic case gives a reasonable limit for the pattern (Fig. 2*B*, dashed line), but it is better described by the linear case (Fig. 2*B*, solid line).

Analysis of the experimental deviations in Fig. 2*B* shows that they significantly follow Eq. **1** for a wiring cost previously used in deterministic optimization theory (18, 21). Another use of Fig. 2*B* could be to find the wiring cost exponent best fitting the experimental pattern. However, this use of Fig. 2*B* would face important limitations. For example, Fig. 2*B* picks out ω as a relevant parameter to study deviations, but its relevance is only valid close to wiring cost exponent 2 (*SI Appendix*). Also, a more quantitative use of Eq. **1** requires taking into account the full distribution of the data and not the simple quantities that can be obtained from Fig. 2*B*, like the envelope or mean of the pattern of deviations. Fig. 2*B* thus allows for a robust qualitative picture, but a quantitative fitting approach must use Eq. **1** and the full data. For this reason, we built a Bayesian estimator that finds the parameters of the objective function that best fit the data to the probability in Eq. **1**. Whereas all of the other procedures presented in this article have predictive value and only use the shape of the objective function, the Bayesian estimator is a fitting procedure that uses extra parameters to describe function f. In order not to make any assumptions about function f, the Bayesian estimator takes into account a wide range of functions, only limited by computational time. A complete description of the Bayesian estimator can be found in *SI Appendix*.

We have used the Bayesian estimator for the wiring cost for *C. elegans*, which has three parameters: α, β and the cost exponent ξ (18, 21) (see *Materials and Methods*). α and β are parameters that weight differently the connections of sensory neurons, interneurons, and motorneurons. There is already an anatomical reason for these weights. We need to transform the connectivity data expressed as number of synapses between any two neurons into number of wires by taking into account the actual anatomy of *C. elegans*. Actual neurites that connect two

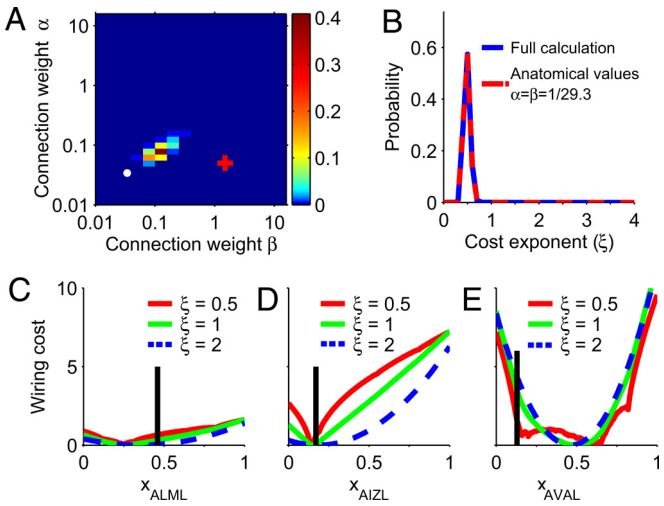


Fig. 3. Bayesian estimation of parameters in wiring cost function of *C. elegans*. (*A*) Probability of α (relative weight for neuron–neuron connections) and β (relative weight for neuron–muscle connections), according to the Bayesian estimator. Most probable values are $\alpha = 0.08$, $\beta = 0.13$. These values are closer to the ones based on *C. elegans* anatomy, $\alpha = \beta = 1/29.3$ (white dot), than to the ones fitting best the data by using deterministic wiring minimization, $\alpha = 0.05$, $\beta = 1.5$ (red plus sign). (*B*) Probability for the cost exponent ξ. Most probable cost exponent is $\xi = 0.49 \pm 0.07$. Results are identical using the complete Bayesian estimation taking into account all values of α, β (blue) as when fixing α, β to their anatomically based values, $\alpha = \beta = 1/29.3$ (red). (*C–E*) Wiring cost along the direction of the position of neurons ALML, AIZL, and AVAL with all other neurons fixed in their experimental positions, for wiring cost exponents $\xi = 0.5$ (red), $\xi = 1.0$ (green), and $\xi = 2.0$ (blue). Black vertical bars: actual soma position. AVAL is far from its optimal position but sits close to a local minimum. The same happens for AVAR (data not shown). DA6 does not improve significantly with the new parameters.

neurons or a neuron and a muscle hold on average 29.3 synapses, whereas neuron-sensor wires only hold one synapse. This anatomical difference may be taken into account by making $\alpha = \beta = 1/29.3$ (ref. 18; *Materials and Methods*). However, so far we have used $\alpha = 0.05$ and $\beta = 1.5$, which are known to best fit the *C. elegans* data within deterministic optimization theory (21). We have used these values to show that objective functions successful in optimization studies can be used to obtain deviations from optimality. For these values, the Bayesian estimator gives the most probable wiring cost exponent of $\xi = 1.3 \pm 0.09$, consistent with our preliminary analysis in Fig. 2*B*. Also, the significance of the pattern for exponent 1.3 is $P = 4{\cdot}10^{-7}$, much better than for the case of quadratic cost.

However, the α and β that are best for deterministic optimization do not need to be the best from the probabilistic point of view. Therefore, we used the Bayesian estimator to find simultaneously the α, β, and ξ that best fit the data to Eq. **1**. We obtain that the most probable α and β (Fig. 3*A*) are close to the anatomically based values, $\alpha = \beta = 1/29.3$ (18) (Fig. 3*A*, white dot). The most probable cost exponent is $\xi = 0.49 \pm 0.07$(SD) (Fig. 3*B*, blue line). Also, fixing α and β to the anatomically based values, the Bayesian estimator finds the same wiring cost exponent $\xi = 0.49 \pm 0.07$ (SD) (Fig. 3*B*, red dashed line). In fact, a cost exponent of $\xi \approx 0.5$ is the most probable for all but the least probable α and β values, *SI Appendix*. Furthermore, by using these estimated parameters, there is an increase of significance, $P \approx 10^{-23}$. Similarly to the cases of linear and quadratic cost exponents, for $\xi \approx 0.5$ the most deviated neurons correspond to directions of flatter wiring cost, Fig. 3 *C* and *D*. A more specific feature of wiring cost with $\xi < 1$ is that it has local minima in the direction of some neurons. This explains the position of neurons like AVAL and AVAR that are now close

to a local minimum (Fig. 3*E*) and for which linear and quadratic wiring costs gave no explanation (Fig. 2*B*). However, the improvement is not only for these special neurons. To show this, we checked that we obtained again the same cost exponent, $\xi = 0.49 \pm 0.06$, using in the Bayesian method all neurons except the three outliers in Fig. 2*B*, DA6, AVAL, and AVAR.

Deviations from Optimality in the Metabolism of *E. coli*. As a second system to test whether Eq. **1** can obtain the suboptimal structure of biological systems, we chose the metabolism of bacterium *E. coli*. It consists of a network of chemical reactions that transform substances in the medium (glucose, oxygen, etc.) into those needed for the bacterium (ATP, amino acids, etc.). Reconstructions of this network are available from biochemical and genetic studies (8–13) (*Materials and Methods*). Thus, we know which reactions take part in the bacterium's metabolism and their stoichiometry. The metabolism is also characterized by the reaction fluxes (or rates). For a given reaction, its flux is the number of molecules of reactants, weighted by their stoichiometric coefficients, that are transformed into products each hour, per gram of bacterial culture (dry weight). These fluxes are controlled by the amount of enzymes of each type produced by the bacterium and have been proposed to be tuned to maximize production of biomass for bacterial growth while producing just enough ATP for maintenance of bacterial functions (8–13). The set of fluxes that are consistent with the stoichiometry of the reactions and maximize biomass production can be computed by linear programming (*Materials and Methods*). The fluxes predicted in this way are in general consistent with experimental results, Fig. 4*A*. It has been proposed that a better predictor of reaction fluxes may be the maximization of ATP production while maintaining a certain growth rate (13). Under the circumstances considered here, maximization of ATP gives a prediction (Fig. 4*B*) very similar to that of maximization of biomass (Fig. 4*A*). This is reasonable, because ATP is one of the most important components of biomass (see *Materials and Methods*), so maximizing biomass implies having a high ATP production. However, although both objectives seem to have their maximum at very similar configurations, we expected the two surfaces to be different enough to predict different deviations from optimality.

Our analysis has two steps. In the first step, we would like to calculate the value of the objective function in the direction of each flux with the rest fixed. But this is not possible because, in general, varying one flux produces a state that is not compatible with the stoichiometry. To avoid this, we allow the other fluxes to change, but to prevent them from taking unrealistic values, we limit them to an interval Δ around the optimum. Then, we fix the flux under study to one of its feasible values and reoptimize the rest of the system within the intervals allowed by Δ. We repeat this procedure for the whole feasible interval to find the upper bound of the objective in the direction of the flux (see *Materials and Methods*). This upper bound is represented for each experimental flux in color scale in Fig. 4*C*. The second step tests whether, as expected from Eq. **1,** experimental data (Fig. 4*C*, white dots) avoids regions where the objective function drops substantially (blue regions in Fig. 4*C*). Fluxes for which the objective stays high only in a small neighborhood of the optimum are expected to have small deviations. Also, fluxes for which the objective drops faster in one direction than in the other are expected to deviate in the direction of slower drop (e.g., flux gltA is expected to have a small negative deviation, Fig. 4*C*). These theoretical results, when using biomass as the objective, have a good correspondence with experimental values (Fig. 4*C*). A significance test was performed by randomly exchanging the deviations among fluxes and calculating the proportion of these configurations with equal or higher average values of the objective, and found $P = 0.0058$ (see *SI Appendix*).

We then tested whether the experimental deviations from the configuration of maximal ATP production can be explained by

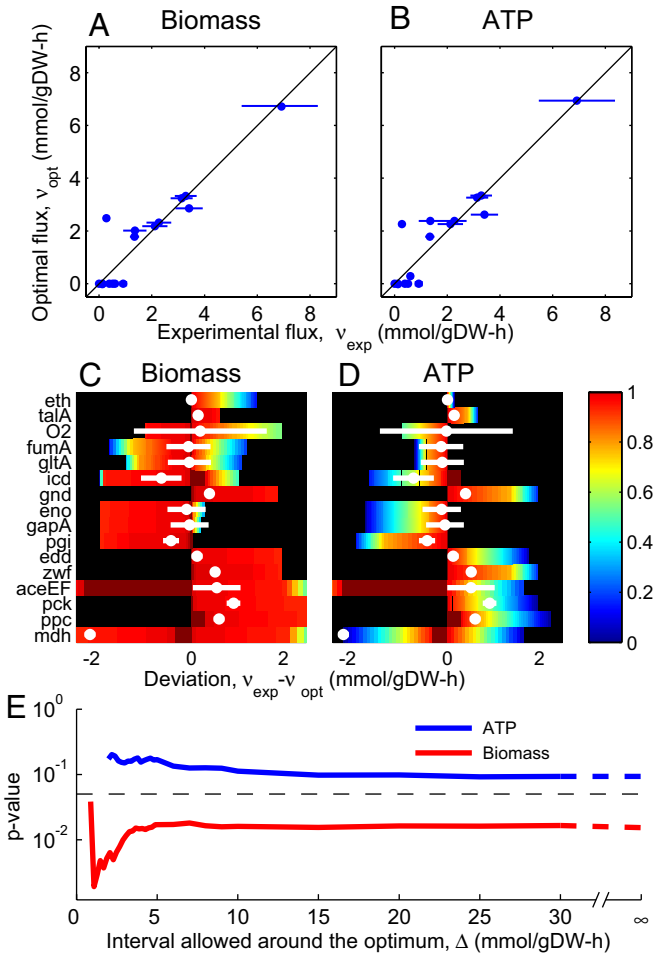


Fig. 4. Deviations from optimality in the metabolic network of *E. coli*. (*A*) Optimal fluxes for maximization of biomass production versus experimental fluxes. Perfect correspondence of deterministic optimization and data would fall in the diagonal. Bars give experimental error reported in ref. 37. (*B*) Optimal fluxes for maximization of ATP production versus experimental fluxes. (*C*) Theoretical and experimental deviations from the optimum for biomass production as objective. For each flux, colors show the value of the objective function (relative to the optimum) as a function of the deviation of the flux. Dark red is reserved for a value of objective of exactly 1, so that maxima are clearly seen. Eq. **1** implies that the fluxes should be at the red regions. White dots are located at the experimental deviations. (*D*) Same as *C* for ATP production as objective. (*E*) *P* value from significance analysis for all possible values of Δ. Red line: significance of the theoretical results using biomass production as objective. Blue line: using ATP production as objective. Black dashed line: $P = 0.05$ significance line.

Eq. **1**. In this case, experimental deviations were found not to be consistent with the theory (Fig. 4*D*, see, e.g., flux mdh, which is predicted to be near the optimum but is, in reality, very far away). The significance test fails in this case, $P = 0.17$. The results presented in Fig. 4 *C* and *D* were obtained for $\Delta = 1.99$ mmol/g dry wt per h, which is the minimum value for which calculations can be performed both for biomass and ATP as objective functions. However, it turns out that our main conclusions do not depend on the value of Δ. For any value of Δ, our theoretical results correspond to experimental data when biomass is the objective function ($P < 0.04$), and not when ATP is the objective function ($P > 0.09$), Fig. 4*E*. All of the results presented here correspond to experiments where bacteria have a growth rate 0.1 h^{-1} (meaning that the bacterial population increases 10% every hour). We further tested that we can obtain the structure of experimental deviations for growth rates of 0.2, 0.3, and 0.4 h^{-1}

(*SI Appendix*). Results for ATP maximization improve slightly for higher growth rates, but biomass typically also improves and it is always the objective most consistent with the experimental data for any value of Δ.

Discussion

We have tested an approach that explains or predicts the statistical structure of deviations from optimality. Using this approach, we have shown that wiring economy and biomass production can explain the experimental data better than previously expected using deterministic optimization theory. Also, the good agreement with the experimental data suggests that the main structure of deviations from optimality need no extra constraints in the objective function. However, we cannot discard the possibility that further refinement of the objective functions may explain finer details of the experimental data. We also gave a method to calculate how significantly the data follow the theoretical prediction, resulting in a P value. This method does not rely on the extraction of some relevant parameter (ω in the case of *C. elegans*) and it is thus generally applicable.

The correspondence between deviations from the optimum and the objective function allowed for a better choice of objective function. In the case of *C. elegans*, whereas wiring cost exponents of values 1 and 2 could correctly obtain the structure of deviations from the optimum, sublinear cost is significantly better. In contrast, previous neuroanatomical studies assumed a wiring cost exponent of a value ≥1 (14–21). At present we do not know which physical processes are contributing most to the wiring cost (i.e., building, maintenance, attenuation, or intracellular transport) to be able to build a realistic model of the origin of a sublinear wiring cost. However, we note that the costs of other better known production or transportation systems, like human-made ones, are in most cases sublinear (an effect that is termed economy of scale for production systems and economy of distance for transportation systems) (38–42).

We believe that the procedures described in this article will be widely applicable to any system for which an optimization principle has been proposed. They allow one to find traces of optimization even in systems far from the optimum and to better select among objective functions. Also, it is possible to use them in systems trapped in a local maximum by studying the shape of the objective function in the neighborhood of that point.

Materials and Methods

Wiring Cost in *C. elegans*. An almost complete reconstruction of the nervous system of *C. elegans* has been obtained from electron microscopy photographs (18, 34, 35). Thus, we know the network connectivity, the soma position for the $n = 279$ neurons and the $S = M = 48$ positions of contact between neurites and muscles or sensors. We used a 1D model where each neuron, sensor, or muscle is represented by a point located in the projection of its 3D position on the anteroposterior axis (18). Data are available at www.wormatlas.org. The total wiring cost for *C. elegans* is (18, 21)

$$W = \frac{1}{2}\alpha \sum_{i,j=1}^{N} A_{ij}|x_i - x_j|^{\xi} + \sum_{i,k=1}^{N,S} B_{ik}|x_i - s_k|^{\xi} + \beta \sum_{i,l=1}^{N,M} C_{il}|x_i - m_l|^{\xi}, \quad [3]$$

where x_i, s_k, and m_l are the positions of neurons, sensors, and muscles, respectively, and **A**, **B**, and **C** are neuron–neuron, neuron–sensor, and neuron–muscle connectivity matrices, respectively. α and β are parameters to take into account differences in average cost of the three kinds of connections. In the actual network, neurites that connect two neurons or a neuron and a muscle hold, on average, 29.3 synapses, whereas neuron–sensor wires only hold one synapse. This anatomical difference may be taken into account by making $\alpha = \beta = 1/29.3$ (18). However, optimization theory has shown a better correspondence with experiments for $\alpha = 0.05$ and $\beta = 1.15$ (21). ξ is a nonnegative exponent, previously argued to be ≈2 (17). Therefore, we use $\xi = 2$, $\alpha = 0.05$, and $\beta = 1.5$, unless otherwise stated.

Deterministic Optimization of *C. elegans*. For quadratic cost ($\xi = 2$), the optimum can be calculated exactly (17, 18, 21). For other exponents, for which no analytical procedure is available, we discretized the worm in 100 bins, computed the cost for each neuron in each position with all other neurons fixed, and selected the bin with lowest cost. Iteration of this algorithm leads to a stationary state in which each neuron is already in the position identified as optimal in the following iteration. In this situation, by construction all partial derivatives will be zero (neglecting the imprecision due to the binning), so the system will be at a local minimum. As the problem is convex for $\xi > 1$, there is only one local minimum which is also the global minimum. For $\xi \leq 1$, we repeat the optimization at least 100 times, starting from random initial conditions, and choose the one ending at lowest cost.

Stochastic optimization of *C. elegans* was performed by stochastic hill climbing, which consists of adding small random variations to the neurons' positions (we added to each position a random number normally distributed with variance 0.001), and accepting the change only when the new cost plus a random number is lower than the previous cost. We ran 5–10 million iterations, checking that both cost and mean deviation had stabilized, which is necessary to satisfy Eq. **1** in the final configuration (*SI Appendix*).

Significance Analysis in *C. elegans*. The permutation test was performed by using the following algorithm: Build a set of positions by randomly permuting the deviations among the neurons and adding them to the optimal positions. In general, some neurons' positions will be outside the limits of the worm. For each of these neurons, find all other neurons such that interchanging the deviations would result in both neurons inside the worm. Choose randomly one of them, and interchange the deviations. Once all neurons are inside the animal, calculate the cost. We repeated this procedure 10^7 times for the computation of the P values. Estimation of P values $<10^{-7}$ was done by approximating the histogram of permutation's cost by a Gaussian distribution, and calculating its cumulative probability from $-\infty$ to the experimental value.

Stoichiometric Model for the Metabolism of *E. coli*. We used a previously published reaction network (13) that, after removal of redundant reactions, is formed by 88 reactions and 60 metabolites. Bacterial growth is incorporated as a reaction consisting in the consumption of metabolites in the proportions measured experimentally in bacterial biomass, and production of the virtual metabolite "Biomass" plus a small amount of subproducts (9) (see Dataset S1 for the meaning of the metabolites' abbreviations),

$$\begin{aligned} 40.2\ \mathrm{ATP} + 0.33\ \mathrm{G6P} + 0.07\ \mathrm{F6P} + 0.12\ \mathrm{GA3P} \\ + 0.86\ \mathrm{3PG} + 0.77\ \mathrm{PEP} + 2.94\ \mathrm{PYR} + 2.41\ \mathrm{ACCOA} \\ + 15.7\ \mathrm{NADPH} + 1.65\ \mathrm{OA} + 0.96\ \mathrm{R5P} + 0.36\ \mathrm{E4P} \\ + 1.28\ \mathrm{AKG} \rightarrow 3\ \mathrm{NADH} + 2.41\ \mathrm{COA} + \mathrm{Biomass}. \end{aligned}$$

This is the reaction whose flux is maximized when biomass is the objective function. When ATP is the objective function, we maximize a reaction which just consists of consumption of ATP. The model can be found in Dataset S1, and is described in detail in *SI Appendix*. We performed linear optimization with the open-source GLPK optimization package (GNU Linear Programming kit; www.gnu.org/software/glpk/), and routines written in Matlab (MathWorks). Our routines are based on COBRA toolbox (36).

Calculation of Maintenance Requirements of *E. coli*. Maintenance requirements are modeled as ATP consumption, with two different contributions: Non-growth-associated maintenance (NGAM), and growth associated maintenance (GAM) (9). These two parameters are computed by a fit to the experimental data, making the growth predicted by the model match the experimental growth (9), see *SI Appendix*. We obtained NGAM = 12.5 mmol of ATP per g dry wt (gDW) per h and GAM = 126 mmol of ATP per gDW. Maintenance is added to the model only when biomass production is the objective function. When ATP is the objective function the network anyway produces as much ATP as possible, and the addition of maintenance requirements makes no difference.

Extra Constraints When ATP Is the Objective Function. When ATP is the objective function, all of the resources of the metabolic network are diverted to ATP production, resulting in zero growth rate. To avoid this unrealistic situation,

when ATP is the objective the growth rate is fixed to its experimental value (13) (see bounds in Dataset S1).

Simulation of Experimental Conditions of *E. coli*. Experimental conditions are simulated by fixing some of the secretion/uptake rates (external fluxes). In all cases glucose uptake rate was fixed to the experimental value. We allowed unlimited uptake of other metabolites present in the medium (CO_2, O_2, and NO_3), and unlimited secretion of the rest of external metabolites (see Dataset S1).

Experimental Data of *E. Coli* Were Taken from Ref. 37. They correspond to ^{13}C labeling experiments on chemostat cultures of *E. coli* (MG1655 strain). A total of eight experiments are presented, in four groups: three experiments with growth rate $\approx 0.1\ h^{-1}$, one experiment with growth rate $\approx 0.2\ h^{-1}$, two with growth rate $\approx 0.3\ h^{-1}$, and two with growth rate $\approx 0.4\ h^{-1}$. For the cases with more than one experiment for the same growth rate, we use the average of the results. These data can be found in Dataset S1.

Calculation of the Upper Bound of the Objective Function in the Direction of Each Flux. First, we constrain to an interval Δ around their optimum all fluxes, except the one under study, the objective and those fixed previously (such as glucose uptake rate, fixed in its experimental value). Then, we find the maximum value of the objective compatible with these constrains for each value of the flux under study. This value is represented in color scale in Fig. 4 *C* and *D*. The feasible interval for the flux under study is larger the higher is Δ. We must select Δ high enough to allow the study of all fluxes with experimental data up to their experimental value. In the case of dilution rate 0.1 h^{-1}, Δ_{min} = 0.88 mmol/gDW·h when biomass is the objective function, and Δ_{min} = 1.99 mmol/gDW·h when ATP is the objective function. In the limit $\Delta \to \infty$, this calculation is the same as robustness analysis (8, 36). See *SI Appendix* for further details.

ACKNOWLEDGMENTS. We are most grateful to Brian Burton and two anonymous referees for critical readings of the manuscript. We also acknowledge discussions with S. Arba-Mosquera, S. Arganda, D. Chklovskii, A. Escudero Berián, A. Koulakov, S. B. Laughlin, V. Pérez Díaz, A. Pérez Escudero, R. Schütz, T. Sharpee, and A. Sorribes. This work was supported by the Ministerio de Ciencia, Tecnología e Innovación (Spain) and the Biociencia program from the Comunidad Autonoma de Madrid (Spain) (G.G.d.P.). A.P-E. and M.R-A. acknowledge fellowships from Ministerio de Ciencia, Tecnología e Innovación (Spain).

1. Barton NH, Briggs DEG, Eisen JA, Goldstein DB, Patel NH (2007) *Evolution* (Cold Spring Harbor Lab Press, Cold Spring Harbor, NY).
2. Alexander RM (1996) *Optima for Animals* (Princeton Univ Press, Princeton).
3. Parker GA, Maynard Smith J (1990) Optimality theory in evolutionary biology. *Nature* 348:27–33.
4. Freeland SJ, Hurst LD (1998) The genetic code is one in a million. *J Mol Evol* 47:238–248.
5. Itzkovitz S, Alon U (2007) The genetic code is nearly optimal for allowing additional information within protein-coding sequences. *Genome Res* 17:405–412.
6. Dekel E, Alon U (2005) Optimality and evolutionary tuning of the expression level of a protein. *Nature* 430:588–592.
7. Tkacik G, Callan CG, Jr, Bialek W (2008) Information flow and optimization in transcriptional regulation. *Proc Natl Acad Sci USA* 105:12265–12270.
8. Palsson BO (2006) *Systems Biology: Properties of Reconstructed Networks* (Cambridge Univ Press, Cambridge, UK).
9. Varma A, Palsson BO (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3010. *Appl Environ Microbiol* 60:3724–3730.
10. Edwards JS, Ibarra RU, Palsson BO (2001) In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 19:125–130.
11. Ibarra RU, Edwards JS, Palsson BO (2002) *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 420:186–189.
12. Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA* 99:15112–15117.
13. Schuetz R, Kuepfer L, Sauer U (2007) Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol* 3:119–133.
14. Chklovskii DB, Schikorski T, Stevens CF (2002) Wiring optimization in cortical circuits. *Neuron* 34:341–347.
15. Klyachko VA, Stevens CF (2003) Connectivity optimization and the positioning of cortical areas. *Proc Natl Acad Sci USA* 100:7937–7941.
16. Buzsaki G, Geiler C, Henze DA, Wang X-J (2004) Interneuron diversity series: Circuit complexity and axon wiring economy of cortical circuits. *Trends Neurosci* 27:186–193.
17. Chklovskii DB (2004) Exact solution for the optimal neuronal layout problem. *Neural Comput* 16:2067–2078.
18. Chen BL, Hall DH, Chklovskii DB (2006) Wiring optimization can relate neuronal structure and function. *Proc Natl Acad Sci USA* 103:4723–4728.
19. Ahn Y-Y, Jeong H, Kim BJ, (2006) Wiring cost in the organization of a biological neuronal network. *Physica A* 367:430–537.
20. Kaiser M, Hiltetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PloS Comp Biol* 2:e95.
21. Perez-Escudero A, de Polavieja G (2007) Optimally wired subnetwork determines neuronanatomy of *Caenorhabditis elegans*. *Proc Natl Acad Sci USA* 104:17180–17185.
22. Laughlin SB (1981) A simple coding procedure enhances a neuron's information capacity. *Z Naturforsch* 36c:910–912.
23. Rieke F, Warland D, Bialek W (1993) Coding efficiency and information rates in sensory neurons. *Europhys Lett* 22:151–156.
24. de Polavieja GG (2002) Errors drive the evolution of biological signalling to costly codes. *J Theor Biol* 214:657–664.
25. Oaten A (1977) Optimal foraging in patches: a case for stochasticity. *Theor Pop Biol* 12: 263–285 '
26. Murray CD (1926) The physiological principle of minimum work. I. The vascular system and the cost of blood volume. *Proc Natl Acad Sci USA* 12:207–214.
27. Murray CD (1926) The physiological principle of minimum work applied to the angle branching of arteries. *J Gen Physiol* 9:235–841.
28. Weibel ER, Gomez DM (1962) Architecture of the human lung. *Science* 137:577–585.
29. Sherman TF (1981) On connecting large vessels to small. *J Gen Physiol* 78:431–453.
30. Gould SJ, Lewontin RC (1979) The spandrels of San Marco and the panglosian paradigm: A critique of the adaptationist program. *Proc R Soc London Ser B* 205:581–598.
31. Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits* (Sinauer, Sunderland, MA).
32. Michalewicz Z, Fogel DB (2000) *How to Solve It: Modern Heuristics* (Springer, New York).
33. Kirkpatrick S, Gelatt CD, Vecchi MP (1983) Optimization by simulated annealing. *Science* 220:671–680.
34. White JG, Southgate E, Thomson JN, Brenner S (1986) The structure of the nervous system of the nematode *Caenorhabditis elegans*. *Philos Trans R Soc London Ser B* 314:1–340.
35. Hall DH, Russell RL (1991) The posterior nervous system of the nematode *Caenorhabditis elegans*: Serial reconstruction of identified neurons and complete pattern of synaptic interactions. *J Neurosci* 11:1–22.
36. Becker SA, et al. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox. *Nat Protocols* 2:727–738.
37. Nanchen A, Schicker A, Sauer U (2006) Nonlinear dependency of intracellular fluxes on growth rate in miniaturized continuous cultures of *Escherichia coli*. *Appl Environ Microbiol* 72:1164–1172.
38. Haldi J, Whitcomb D (1967) Economies of scale in industrial plants. *J Polit Econ* 75:373–385.
39. Nerlove M (1963) Returns to scale in electricity supply. *Measurement in Economics*, eds Christ F, et al (Stanford Univ Press, Stanford, CA), pp 167–198.
40. Christensen LR, Greene WH (1976) Economies of scale in U.S. electric power generation. *J Polit* 4:655–676.
41. Janic M (2007) Modelling the full costs of an intermodal and road freight transport network. *Transportation Res D* 12:33–44.
42. McCann P (2001) A proof of the relationship between optimal vehicle size, haulage length and structure of distance-transport costs. *Transportation Res A* 35:671–693.

Supporting Appendix

Alfonso Pérez-Escudero, Marta Rivera-Alba,

Gonzalo G. de Polavieja

September 1, 2009

Contents

1 Derivation of $P(\vec{x}) = f(Z(\vec{x}))$ (Equation 1 in main text) from simple models

1.1 Derivation for the case of stochastic optimization

We consider an optimizer that can explore the whole space of states. This optimizer does not evaluate exactly the objective function $Z(\vec{x})$ (where $\vec{x}$ denotes the state). Rather, it evaluates a stochastic objective function, $Z_s(\vec{x}) = Z(\vec{x}) + \eta$, where η is a random variable independent of $\vec{x}$ and governed by the probability distribution $\rho(\eta)$.

The probability that the stochastic objective function at point $\vec{x}_i$ is higher than at point $\vec{x}_j$ is

$$P\left(Z_s(\vec{x}_i) > Z_s(\vec{x}_j)\right) = \\ P\left(Z(\vec{x}_i) + \eta_i > Z(\vec{x}_j) + \eta_j\right) = P\left(Z(\vec{x}_i) - Z(\vec{x}_j) > \eta_j - \eta_i\right).$$

As the subtraction of two random variables is another random variable, there exists a random variable r governed by the probability distribution $\rho_r(r)$, so that

$$P\left(Z(\vec{x}_i) - Z(\vec{x}_j) > \eta_j - \eta_i\right) = P\left(Z(\vec{x}_i) - Z(\vec{x}_k) > r\right).$$

Therefore,

$$P\left(Z_s(\vec{x}_i) > Z_s(\vec{x}_j)\right) = P\left(Z(\vec{x}_i) - Z(\vec{x}_j) > r\right) = \int_{-\infty}^{Z(\vec{x}_i)-Z(\vec{x}_j)} \rho_r(r)\, dr = R\left(Z(\vec{x}_i) - Z(\vec{x}_j)\right),$$

where function R is the cumulative probability of distribution ρ_r. Note that the probability depends only on the difference of the objective function at the two points.

Now, the probability $P(\vec{x}_i)$ of identifying $\vec{x}_i$ as the optimum is equal to the probability that the stochastic objective function at point $\vec{x}_i$ is higher than the stochastic objective function at any other point of the space of states, that is, $P\left(Z_s(\vec{x}_i) > Z_s(\vec{x}_j)\ \forall j \neq i\right)$. At this point we assume that the space of states is

discretized, with index j running over all possible states (note that this discretization may be as tight as necessary to mimic the continuous case). As the random component is independent for each state, the probability for all points is just the product of the probabilities for each point,

$$P\left(\vec{x}_i\right) = P\left(Z_s(\vec{x}_i) > Z_s(\vec{x}_j)\ \forall j \neq i\right) = \frac{\prod_{j\neq i} R\left(Z(\vec{x}_i) - Z(\vec{x}_j)\right)}{\sum_k \left[\prod_{j\neq k} R\left(Z(\vec{x}_i) - Z(\vec{x}_j)\right)\right]}, \tag{S1}$$

where the denominator is just a normalization term, equal for all points. This probability is, in general, very difficult to compute. However, note that the only dependence on $\vec{x}_i$ is through $Z(\vec{x}_i)$. Therefore, all points with the same value of the objective function will have the same probability, meaning that the probability can be written as a function of the objective function, $P(\vec{x}) = f_\rho(Z(\vec{x}))$. □

Figure S1A shows tests of this result for 2-dimensional cost (linear, quadratic and cubic) with stochastic components following normal, uniform, exponential and beta distributions. **Figure S1B** shows a test for the case of stochastic hill climbing.

1.2 Derivation for the case of finite-time optimization

We consider an optimizer that explores the space of states by randomly jumping from one state to the other, as a simplified model of mutations in biological systems. We assume that jumps are equally likely to any state, so the jumping probability distribution is a constant. Therefore, the probability of jumping to state $\vec{x}$ in the m-th iteration is

$$\rho_{\text{jump}}\left(\vec{x}_m\right) = \frac{1}{V_{\text{tot}}}, \tag{S2}$$

where V_{tot} is the total hypervolume of the space of states and acts as normalization constant. We denote $\vec{x}_m$ the state resulting from the random jump in the m-th iteration. After a jump, the new state is accepted if the new value of the objective function Z is higher than the original one (note that the optimization scheme is deterministic, in the sense that no noise is added to the objective function). After K succesive jumps, the system will be in point $\vec{x}$ if at least one jump led to $\vec{x}$ and all other jumps led to points with lower values of the objective function. This probability is

$$P_K\left(\vec{x}\right) = \sum_{m=1}^{K} \left(\underbrace{\rho_{\text{jump}}\left(\vec{x}_m = \vec{x}\right)}_{\substack{\text{Prob. that} \\ \text{jump } m \\ \text{leads to } \vec{x}}} \underbrace{\prod_{\substack{m'=1 \\ m' \neq m}}^{K} \underbrace{\int_{Z(\vec{x}_{m'}) < Z(\vec{x})} \rho_{\text{jump}}\left(\vec{x}_{m'}\right)\, d\vec{x}_{m'}}_{\substack{\text{Prob. that jump } m' \text{ leads} \\ \text{to a value of } Z \text{ lower than } Z(\vec{x})}}}_{\substack{\text{Prob. that all jumps except the } m\text{-th} \\ \text{lead to a value of } Z \text{ lower than } Z(\vec{x})}} \right),$$

where the integral is a multiple one over the portion of the space where the objective function is lower than $Z\left(\vec{x}\right)$. The probability of jumping comes out of the integral because it is constant for all $\vec{x}$, Equation (S2). Therefore, we get

$$P_K\left(\vec{x}\right) = \sum_{m=1}^{K} \left(\frac{1}{V_{\text{tot}}} \prod_{\substack{m'=1 \\ m' \neq m}}^{K} \frac{1}{V_{\text{tot}}} \int_{Z(\vec{x}') < Z(\vec{x})} d\vec{x}' \right).$$

Terms depend neither on m nor m'. This means that the order of jumps does not matter, a consequence of the fact that each jump is independent of the past history.

Mathematically, this allows to replace the summation for a multiplication by the number of terms, and the product for a power to the number of terms, giving

$$P_K(\vec{x}) = \frac{K}{V_{\text{tot}}^K} \left(\int_{Z(\vec{x}')<Z(\vec{x})} d\vec{x}' \right)^{K-1} .$$

Note that the integral is the hypervolume where the objective function is lower than $Z(\vec{x})$. This quantity is usually difficult to compute. However, we know that it is equal for all states $\vec{x}$ for which the objective function $Z(\vec{x})$ is the same. Therefore, we can express this hypervolume as a function $V(Z(\vec{x}))$, that depends only on $Z(\vec{x})$. So the probability is finally

$$P_K(\vec{x}) = \frac{K}{V_{\text{tot}}} \left(\frac{V(Z(\vec{x}))}{V_{\text{tot}}} \right)^{K-1} = f_K(Z(\vec{x})) . \quad \square \tag{S3}$$

Interestingly, in this case the probability is completely determined, depending only on the objective function Z and the number of iterations K.

Figure S1A shows tests of this result for 2-dimensional cost (linear, quadratic and cubic).

1.3 Illustration of stochasticity, finite time and stochastic hill climbing for a two-dimensional system

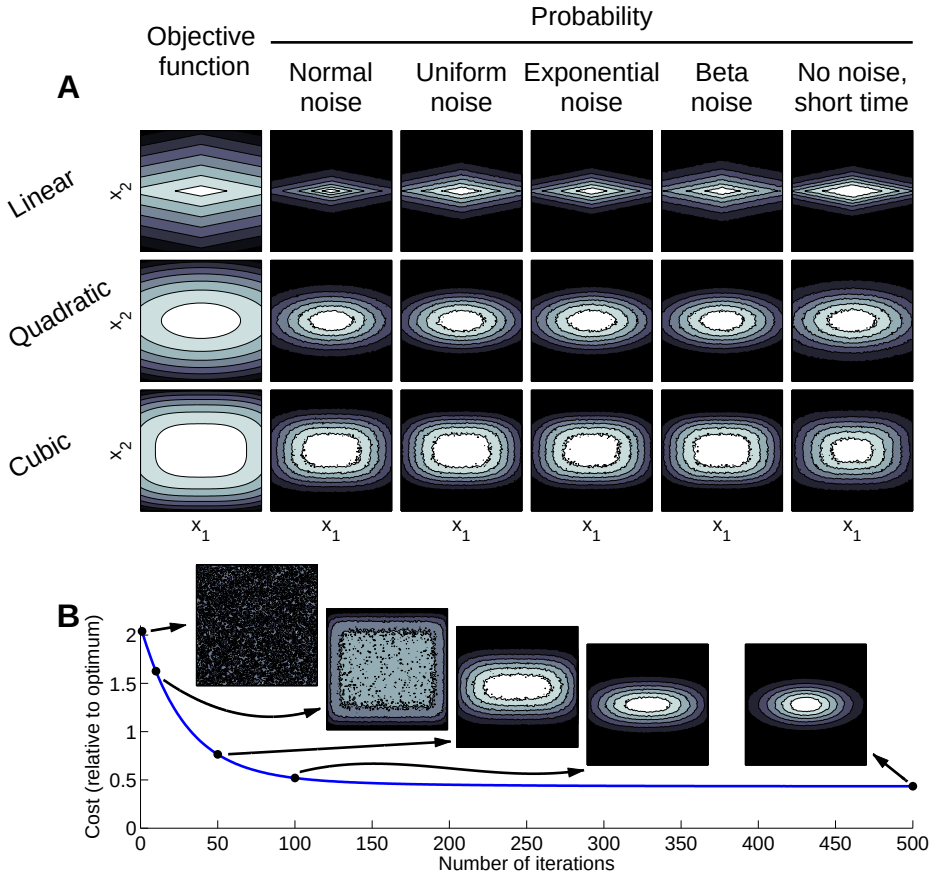


Figure S1: The probability is a function of the objective function. **(A) First column:** Objective function Z for the simple case $Z = |x_1|^{\xi} + 5|x_2|^{\xi}$, and three different exponents ($\xi = 1$, 2 and 3). **Second to fifth columns:** Probability for stochastic optimization (stochastic component taken from normal, uniform, exponential or beta distributions, with the variance adjusted so that final standard deviation in x_1 is about 0.5). **Last column:** Probability after deterministic optimization during insufficient time (K iterations of randomly choosing any point of the space and accepting it if the cost decreases, $K = 6$ for linear cost and $K = 3$ for quadratic and cubic costs). All probabilities are histograms of 10^7 simulations. **(B)** Stochastic hill climbing calculation for the quadratic objective function and stochastic component from normal probability distribution. The probability $P(x_1, x_2)$ converges to a function of the objective. Results obtained from 10^7 simulations.

2 Supporting Information for *Caenorhabditis elegans*

2.1 Quadratic wiring cost for one neuron's position (Equation 2 of main text)

The quadratic wiring cost is

$$W(\vec{x}) = \frac{\alpha}{2} \sum_{i,j=1}^{N} A_{ij} \left(x_i - x_j\right)^2 + \sum_{i,k=1}^{N,S} B_{ik} \left(x_i - s_k\right)^2 + \beta \sum_{i,l=1}^{N,M} C_{il} \left(x_i - m_l\right)^2, \quad \text{(S4)}$$

where x_i, s_k and m_l are the positions of neurons, sensors and muscles, respectively. $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are the connectivity matrices for neuron-neuron, neuron-sensor and neuron-muscle connections, respectively. α and β take into account differences in average cost for the three kinds of connections. Here we will use a compact notation, defining matrix $\mathbf{D}$ as

$$\mathbf{D} = \begin{pmatrix} \alpha\mathbf{A} & \mathbf{B} & \beta\mathbf{C} \\ \mathbf{B}' & \mathbf{0} & \mathbf{0} \\ \beta\mathbf{C}' & \mathbf{0} & \mathbf{0} \end{pmatrix}, \quad \text{(S5)}$$

where primed matrices are transposed. In terms of this matrix, the wiring cost is

$$W(\vec{x}) = \frac{1}{2} \sum_{i,j=1}^{N+M+S} D_{ij} \left(x_i - x_j\right)^2, \quad \text{(S6)}$$

where now x_i stands for neurons' positions for $i = 1 \ldots N$, sensors' positions for $i = N+1 \ldots N+S$ and muscles' positions for $i = N+S+1 \ldots N+S+M$.

We are interested on how the optimization affects neuron p, so we will study the cost in its direction. First, we find the optimal position for neuron p with other neurons fixed, by setting $\frac{\partial W}{\partial x_p} = 0$:

$$\frac{\partial W}{\partial x_p} = \frac{1}{2} \sum_{i,j=1}^{N+M+S} D_{ij} \left[2\delta_{ip}(x_i - x_j) - 2\delta_{jp}(x_i - x_j)\right] = 0.$$

As matrix $\mathbf{D}$ is symmetric, we have that

$$\frac{\partial W}{\partial x_p} = 2\sum_{j=1}^{N+M+S} D_{pj}(x_p - x_j) = 0.$$

We can directly work out the optimal value of x_p, getting

$$x_p^{\text{opt}} = \frac{\sum_{j=1}^{N+M+S} D_{pj} x_j}{\sum_{j=1}^{N+M+S} D_{pj}}. \tag{S7}$$

where the superscript 'opt' stands for 'optimum'.

Now we come back to the cost W (Equation (S6)). We define W_1 as a cost term that does not depend on x_p,

$$W_1 \equiv \frac{1}{2} \sum_{\substack{i,j=1 \\ i,j\neq p}}^{N+M+S} D_{ij}(x_i - x_j)^2.$$

Then, the total cost can be written as

$$W(x_p) = W_1 + \sum_{j=1}^{N+M+S} D_{pj} \left(x_p - x_j\right)^2,$$

where we write $W(x_p)$ to emphasize that we will study the cost in the direction of neuron p, keeping all other neurons fixed. We are interested on the behaviour of W as x_p deviates from the optimum. Let us define $\Delta x_p = x_p - x_p^{\text{opt}}$, and substitute x_p in terms of this deviation:

$$W(x_p) = W_1 + \sum_{j=1}^{N+M+S} D_{pj} \left(x_p^{\text{opt}} + \Delta x_p - x_j\right)^2.$$

By expanding the brackets and separating the terms with some dependence on Δx_p, we get

$$W(x_p) = W_1 + \sum_{j=1}^{N+M+S} D_{pj} \left(x_p^{\text{opt}} - x_j\right)^2 + \Delta x_p^2 \sum_{j=1}^{N+M+S} D_{pj} - 2\Delta x_p \sum_{j=1}^{N+M+S} D_{pj} x_j + 2\Delta x_p x_p^{\text{opt}} \sum_{j=1}^{N+M+S} D_{pj}.$$

When $\Delta x_p = 0$ (so $x_p = x_p^{\text{opt}}$) only the two first terms remain, so we call them $W(x_p^{\text{opt}})$. We substitute the value of x_p^{opt} (Eqn. (S7)), and we get

$$W(x_p) = W(x_p^{\text{opt}}) + \Delta x_p^2 \sum_{j=1}^{N+M+S} D_{pj} - 2\Delta x_p \sum_{j=1}^{N+M+S} D_{pj} x_j + 2\Delta x_p \frac{\sum_{j=1}^{N+M+S} D_{pj} x_j}{\sum_{j=1}^{N+M+S} D_{pj}} \sum_{j=1}^{N+M+S} D_{pj}.$$

The last two terms cancel out, and finally

$$W(x_p) - W(x_p^{\text{opt}}) = \omega_p \Delta x_p^2, \tag{S8}$$

where

$$\begin{aligned} \omega_p = & \sum_{j=1}^{N+M+S} D_{pj} \\ = & \alpha \sum_{j=1}^{N} A_{pj} + \sum_{k=1}^{S} B_{pk} + \beta \sum_{l=1}^{M} C_{pl} \end{aligned}$$

is the effective number of wires of neuron p. □

Note that the above derivation is not restricted to the case when all other neurons are fixed in their optimal positions. They can be fixed in any position (for example, they could be fixed in their experimental positions). Interestingly, the position of the other neurons affects x_p^{opt}, but has no effect on how the cost increases as x_p deviates from it.

2.2 Isoprobability lines for *C. elegans* in the ω vs $|\Delta x|$ plots (lines in Figure 2B,D of main text)

Plotting the effective number of connections ω versus the deviations $|\Delta x|$ may be used for a fast check of our predictions. While the qualitative pattern is very robust and does not depend on the cost parameters, the shape of the pattern does depend on the cost exponent, ξ. In particular, the deviations scale on average as $|\Delta x_i| \propto \omega_i^{-1/\xi}$. In this note we will see why this rule is satisfied approximately in *C. elegans*. We start from Equation (2) of the main text, or Equation S8 of Supporting Text 2.1 for the quadratic cost in the direction of one neuron, with all other neurons fixed, which we repeat here:

$$W(x_i) - W(x_i^{\text{opt}}) = \omega_p \Delta x_p^2, \tag{S9}$$

where $\Delta x_i \equiv x_i - x_i^{\text{opt}}$. Now, if we assume that neurons are independent of each other, we can write the whole cost as

$$W(\vec{x}) = \sum_{i=1}^{N} \omega_i |\Delta x_i|^2, \tag{S10}$$

where we have neglected the constant $W(x_i^{\text{opt}})$ because it has no effect on the derivation presented here. The approximation of independent neurons is reasonable in the case of *C. elegans*, because most of the cost is due to connections to sensors and muscles, which remain fixed. Next, we generalize the expression, assuming that for other exponents ξ (near to 2), the cost is approximately

$$W(\vec{x}) = \sum_{i=1}^{N} \omega_i |\Delta x_i|^{\xi}, \tag{S11}$$

The probability is a function of the cost (Equation (1) of the main text), so

$$P(\vec{x}) = f\left(\sum_{i=1}^{N} \omega_i |\Delta x_i|^{\xi}\right).$$

We make the change of variable

$$u_i = \omega_i^{1/\xi} \Delta x_i.$$

The determinant of the Jacobian matrix of this change is a constant, $J = \prod_{i=1}^{N} \omega_i^{1/\xi}$, so the probability in the new variables is

$$P(\vec{u}) = \frac{1}{J} f\left(\sum_{i=1}^{N} |u_i|^{\xi}\right). \tag{S12}$$

The probability is symmetrical in the u-variables, but the interval in which it is defined is different for each variable: It was already different in the x-space, because each neuron is at a different distance from the extremes of the nematode. In the u-space this is still more evident, because each dimension has shrinked differently, depending on ω_i. However, border effects become negligible if the probability becomes sufficiently small before reaching the border of the nematode. In this case, extending the limits to infinity does not change the results, and the probability becomes truly symmetric.

In these conditions, the symmetry of Equation S12 ensures that any operation we make will give the same result for all neurons. In particular, to obtain the probability distribution for one single neuron we must integrate over all other dimensions. This integral will not be computable in general, but the symmetry ensures that it will be the same for all neurons. Therefore, any calculation such that expected value will give the same result for all neurons, say $\langle |u| \rangle$. Now, we undo the variable change, and we get that for each variable

$$\langle |\Delta x_i| \rangle = \langle |u| \rangle / \omega_i^{1/\xi}. \quad \square$$

Note that this proof is valid not only for the expected value, but also for any other operation on the probability, such as the variance or the point where the cumulative distribution reaches a certain threshold.

To illustrate this in a case where the condition of separable cost (eqn. (S11)) is exactly satisfied, we simulated a system with 200 neurons, with possible positions between -1 and 1, and whose cost is that of equation (S11). We assign the values of ω randomly, taking them from a uniform distribution between 0 and 1. Then, we run simulations for cost exponents $\xi = 1$, $\xi = 2$ and $\xi = 3$. Stochastic optimization

of the system was performed by stochastic hill climbing, with the stochastic component taken from normal, uniform, exponential and beta distributions. In all cases the variance of the stochastic component was adjusted so that mean deviation was around 0.1. In order to have enough statistics, we run 400 simulations for each case. Cumulative density plots for all cases are shown in **Figure S2**, with theoretical lines that follow $\omega \propto |\Delta x|^{-\xi}$. In all cases the simulations agree with the theoretical lines, except the region of low ω in the linear case. This is because in this case the probability is non-negligible beyond the limits of the system, and therefore normalization issues affect the solution, causing the deviation from the theoretical line.

Next, we tested the validity of this result for the *C. elegans* wiring cost. In this case, the condition (S11) is satisfied only approximately. We calculated the cumulative distribution for synthetic datasets of *C. elegans*, calculated by stochastic hill climbing, with cost exponent $\xi = 1$ (**Figure S3A**) and $\xi = 2$ (**Figure S3B**). In both cases, we find a good agreement with the theoretical results.

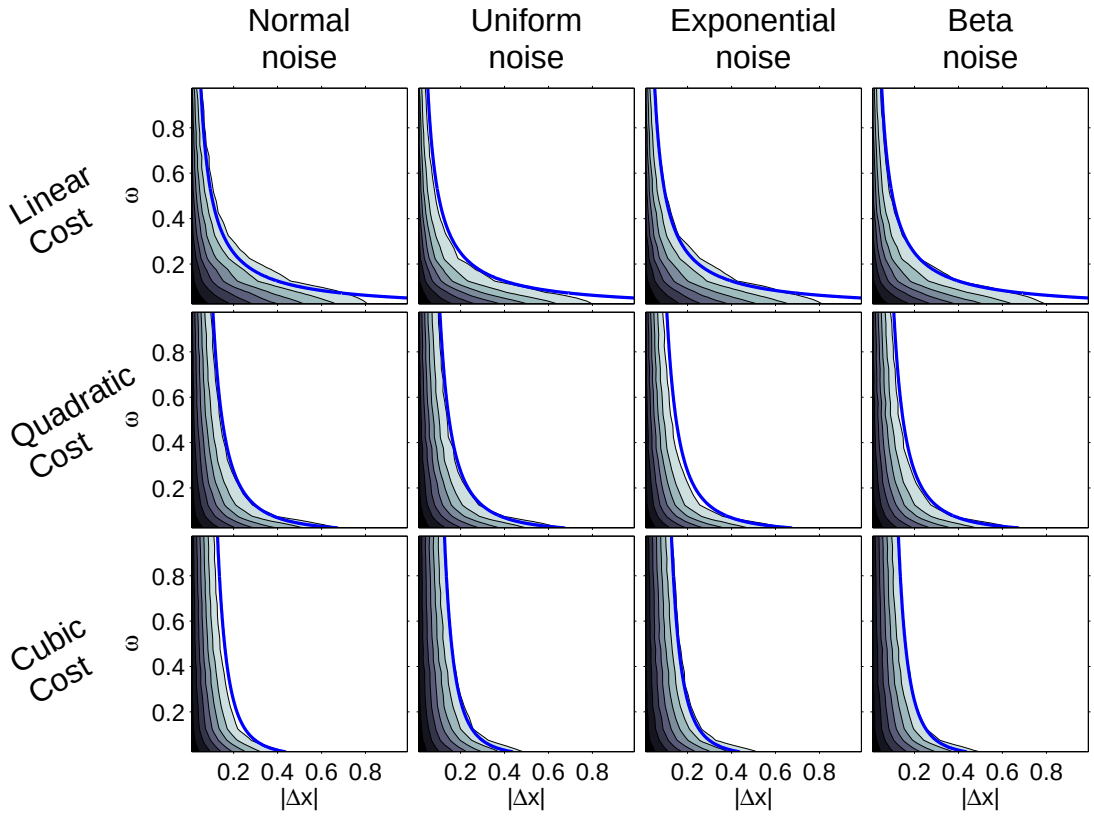


Figure S2: Cumulative probability follows the rule $\omega \propto |\Delta x|^{-\xi}$. Test in the case of independent neurons for linear, quadratic and cubic costs (top, middle and bottom rows, respectively), with different stochastic components added to the cost (following normal, uniform, exponential and beta distributions, shown in first, second, third and fourth columns respectively). Blue lines follow $\omega \propto |\Delta x|^{-\xi}$, with ξ equal to 1, 2 and 3 for the first, second and third rows, respectively.

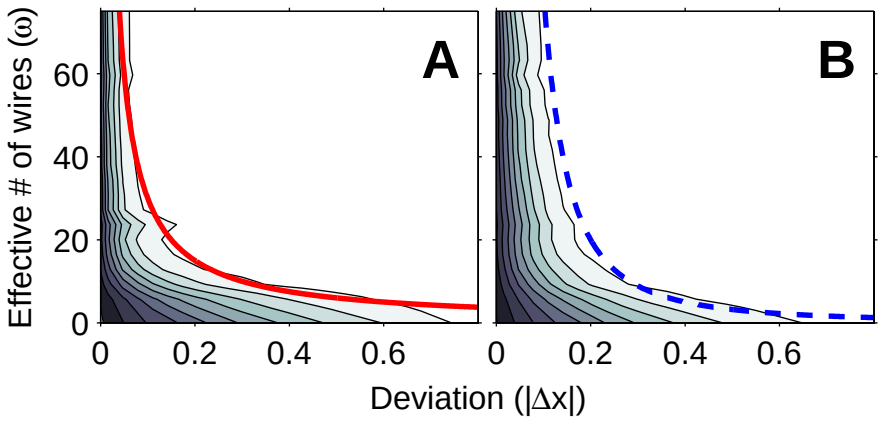


Figure S3: Cumulative probability follows the rule $\omega \propto |\Delta x|^{-\xi}$ for the *C. elegans* wiring cost. **(A)** Cumulative distribution for 380 stochastic optimizations of *C. elegans*, assuming linear cost ($\xi = 1$), with normal stochastic component. Red line follows $\omega \propto |\Delta x|^{-1}$. **(B)** Cumulative distribution for 480 stochastic optimizations of *C. elegans*, assuming quadratic cost ($\xi = 2$), with normal stochastic component. Blue dashed line follows $\omega \propto |\Delta x|^{-2}$.

2.3 Controls for *Caenorhabditis elegans*

2.3.1 Pattern in ω vs $|\Delta x|$ plots is not due to sampling

The experimental pattern shown in **Figure 2** of the main text (reproduced in **Figure S4I**, top), seems to confirm the prediction that neurons with lower ω deviate on average more than neurons with higher ω. However, it might happen that deviations did not depend on ω but, because there are more neurons with low ω, the tail of the distribution is better sampled in this case, originating the visual pattern. Our significance test based on permutations of the deviations (that yielded $p_{\text{perm}} = 0.0007$) already eliminates the possibility that deviations are independent of the shape of the cost surface, but we still need to show that this dependence is correctly captured by parameter ω. For this, we separated the deviations in two groups, those with $\omega \leq 15$ and those with $\omega > 15$, and studied the statistics of the two groups separately (histograms shown in **Figure S4**, second row). First, we used a non-parametric statistical test to find out whether both populations had significantly different standard deviation. This test consists of mixing the two populations, randomizing the order, and then building again two groups with the same number of elements as the original ones. The proportion of these synthetic populations that have standard deviations at least as different as the actual ones is the p-value, and we obtained $p_{\text{non-param}} = 7 \cdot 10^{-5}$. In addition to this test, we built a Bayesian estimator to find the standard deviation that best fits the two populations, assuming that they follow an exponential distribution (see third row of **Figure S4**). From this estimate we computed the p-value as the probability that the population of high ω has a higher or equal standard deviation than the population of low ω, and got $p_{\text{Bayes}} = 0.0003$. **Figure S4II,III** shows the same tests for controls built by adding random noise to the optimal positions. These controls are found non-significant by the three statistical tests. **Figure S4IV** shows the results of stochastic optimization, which passes the three tests.

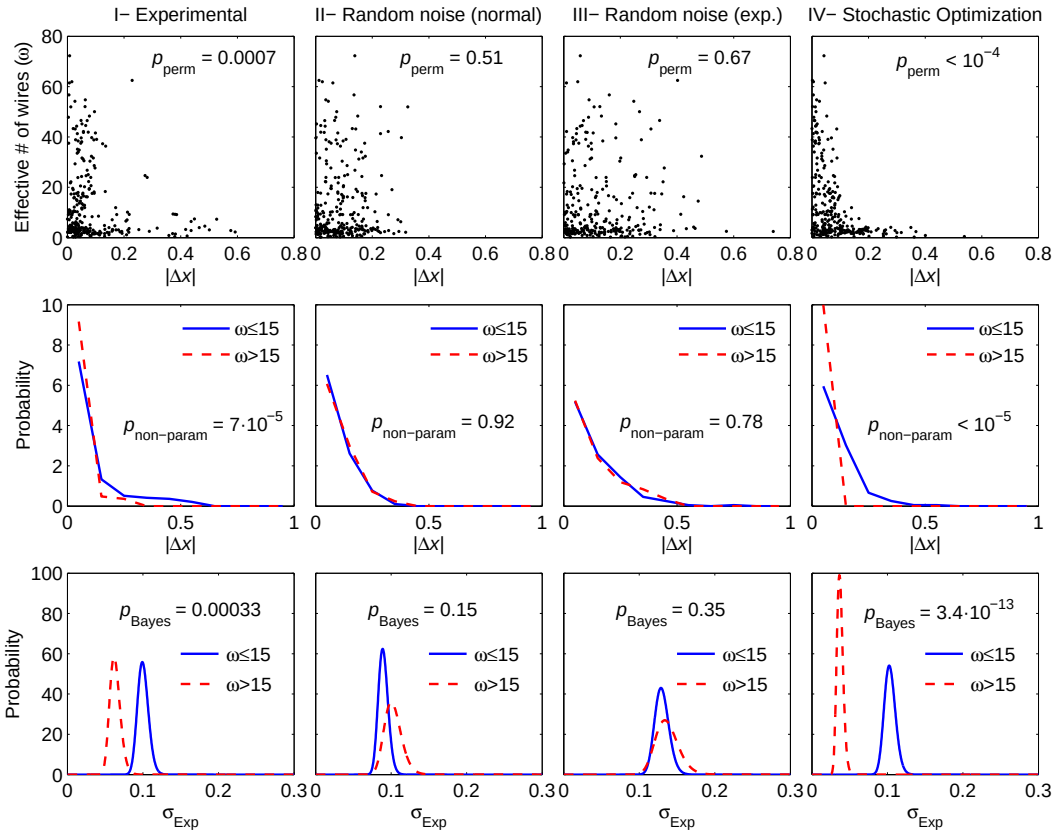


Figure S4: The pattern in the ω vs $|\Delta x|$ plots for *C. elegans* is not due to sampling. **Top row:** Effective number of wires (ω) vs. deviation from optimum. Inset text shows the p-value according to the permutations significance test. **Middle row:** Histograms of deviations for the two groups of neurons (solid blue, $\omega < 15$; dashed red, $\omega > 15$). Inset text shows the p-value for the two variances to be different (non-parametric test). **Bottom row:** Probability for each standard deviation, according to the bayesian estimator (solid blue, $\omega < 15$; dashed red, $\omega > 15$). Inset text shows the p-value for the two variances to be different (bayesian test). **Columns: I-** Experimental data. **II-** Synthetic positions, created as optimal positions + random numbers taken from a normal distribution. **III-** Synthetic positions, created as optimal positions + random numbers taken from an exponential distribution. **IV-** Synthetic positions, created by stochastic optimization (stochastic component taken from a normal distribution).

2.3.2 Deviations are not due to experimental errors in the localization of sensors and muscles

It might happen that incorrect localization of muscles and sensors caused the calculations to yield wrong optimal positions for the neurons. We added random noise (normal and exponential, of several variances up to 0.2) to the positions of muscles and sensors, repeated the optimization and computed the deviations with respect to the original optimization. No pattern in the ω vs $|\Delta x|$ plots resulted from these tests, and it is regarded non-significant by all tests (data not shown).

2.3.3 Deviations are not due to reconstruction errors in the connectivity matrix

Errors in the reconstruction of the connectivity matrix may produce deviations that depend on ω, because the effect of a new connection is weaker if it competes with many other ones to determine the neuron's optimal position. The authors of the reconstruction claim that it is 97% complete, and list the 61 neurons that might be affected by missing parts [1, 2]. First, we checked that removing these 61 neurons the permutation significance test still regards the data as significantly following the predictions of stochastic optimization ($p_{\text{perm}} = 0.002$). If deviations were mainly due to reconstruction errors, one would expect for the partially-reconstructed neurons to deviate more than the fully-reconstructed ones. This is not the case, as shown in (**Figure S5A**). Still, it might be that fully-reconstructed neurons are deviated due to interactions with the partially-reconstructed ones. To check that this is not the case, we simulated the effect of incomplete reconstruction by adding random connections to the 61 partially-reconstructed neurons until their mean deviation equals the mean experimental deviation. We find that fully-reconstructed neurons never deviate more than 5%, much lower than the experimental value (compare **Figure S5A** and **Figure S5B**). Even assuming that all neurons may have missing connections up to the reported 3%, the deviations are much lower than in the ex-

perimental case, **Figure S5C**. Even three times more missing connections do not explain the experimental deviations, **Figure S5D**. We would need to assume about 18% missing connections (and affecting all neurons, even the fully-reconstructed ones) to reproduce the experimental deviations, **Figure S5E**.

One might argue that experimental deviations come from a combination of connectivity errors and errors in the measurement of the neurons and/or sensors and muscles positions. There is no explicit reference about the precision in these measurements, so we estimated it from the available information: Since each photograph covers about 150 nm (about 0.015% of the animal length), soma positions should be very accurate. But positions are given rounded to the 1% of the animal length, so to be conservative we take 2% as the estimate of their error. The region compatible with experimental accuracy (up to 3% connections missing and up to 2% error in positions) is covered by magenta strips in **Figure S5F,G**. We run simulations for each proportion of missing connections and each standard deviation of the normal noise added to the position of neurons (**Fig. S5F**) or organs (**Fig. S5G**). For each simulation we computed the standard deviation of deviations from the optimum (color scale in **Figure S5F,G**), and performed the permutations significance test (dark areas in **Fig. S5F,G** is regarded as non-significant*). The region where the simulations have deviations compatible with the experimental ones (covered by white strips in **Fig. S5F,G**) is far away from the region compatible with experimental accuracy.

*Note that the normal noise added to positions does not correlate with ω, and therefore destroys the pattern when its effect is larger than the effect of missing connections.

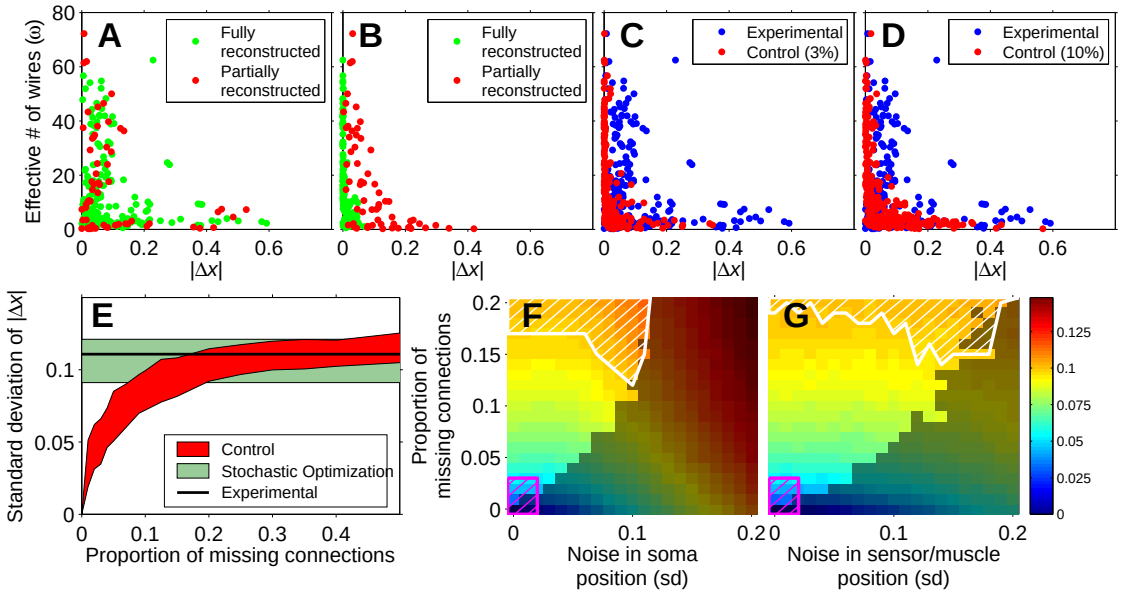


Figure S5: Errors in the connectivity matrix do not explain the experimental deviations. **(A)** Deviations pattern for the experimental case. Partially-reconstructed neurons (red) do not deviate more than fully-reconstructed ones (green). **(B)** Simulation with 20% random connections added to the partially-reconstructed neurons. **(C)** Comparison of experimental deviations (blue) and a simulation with 3% random connections added to all neurons (red). **(D)** Comparison of experimental deviations (blue) and a simulation with 10% random connections added to all neurons (red). **(E)** Red: Standard deviation (sd) of deviations, for simulations with different proportions of random connections added (95% confidence interval). Green: 95% confidence interval for std of deviations resulting from stochastic optimization. Black line: Experimental sd of deviations. **(F)** Standard deviation of deviations, for simulations with different proportions of added connections and different amount of normal noise added to the soma positions. Magenta strips: Region compatible with experimental accuracy. White strips: Region compatible with experimental deviations. Dark region: Region where the permutation significance test results non-significant. **(G)** The same as (F), but the normal noise is added in the positions of sensors and muscles.

2.4 Bayesian estimator for the cost parameters in *C. elegans*

In this section we describe the Bayesian estimator used to select, among a family of possible objective functions, those that better fit the experimental data, taking into account not only the optimal configuration corresponding to each function, but also the shape of the objective function in the neighborhood of the optimum. Thus, we want to find the probability that the experimental configuration ($\vec{x}$) comes from a given set of parameters for the objective function ($\vec{\varphi}$). This probability can be written as $P(\vec{\varphi}|\vec{x})$, that is, the probability for the parameters, provided the observed configuration. In the following section we will show how to extract this probability from the inverse probability, $P(\vec{x}|\vec{\varphi})$, which, by Equation (1) of the main text is a function of the objective Z.

2.4.1 General theory

The wiring cost for *C. elegans* has three parameters, the exponent ξ and the two weights α and β (see Materials and Methods). Our aim in this section is to find an expression for $P(\xi, \alpha, \beta|\vec{x})$, that is, the probability of the cost exponent ξ and the relative weights α and β, given the actual positions $\vec{x}$ of the neurons. We define $\vec{\varphi} \equiv (\xi, \alpha, \beta)$. Thus, we look for $P(\vec{\varphi}|\vec{x})$.

We will make use of the fact that the probability for $\vec{x}$ is a function of the cost (see Supplementary Text 1.1 and 1.2 and Equation (1) of the main text), and ξ, α and β ($\vec{\varphi}$ in our new notation) are parameters of the cost. Therefore, we know that the probability for a configuration $\vec{x}$, given the parameters $\vec{\varphi}$, is

$$P(\vec{x}|\vec{\varphi}) = f(W(\vec{x};\vec{\varphi})).$$

We do not know the form of function f. Therefore, we must take into account all possible options. In principle, it is valid any function that decreases monotonically as W increases, and that is normalized so that $\int f(W(\vec{x};\vec{\varphi}))\, d\vec{x} = 1$. Let us assume

that a set of k parameters, $\vec{\lambda}$, indexes all valid f functions (e.g. polynomial with all possible coefficients, exponential with all possible variances, etc.).[†] Therefore, for a given $f_{\vec{\lambda}}$ we know the probability for $\vec{x}$, which is

$$P(\vec{x}|\vec{\varphi},\vec{\lambda}) = f_\lambda(W(\vec{x};\vec{\varphi})). \tag{S13}$$

According to Bayes' rule,

$$P(\vec{\varphi},\vec{\lambda}|\vec{x}) = \frac{P(\vec{x}|\vec{\varphi},\vec{\lambda})P(\vec{\varphi},\vec{\lambda})}{P(\vec{x})}. \tag{S14}$$

The denominator $P(\vec{x})$ is a normalization factor that can be computed by integrating the numerator over all possible cases, so Bayes' rule becomes

$$P(\vec{\varphi},\vec{\lambda}|\vec{x}) = \frac{P(\vec{x}|\vec{\varphi},\vec{\lambda})P(\vec{\varphi},\vec{\lambda})}{\iint P(\vec{x}|\vec{\varphi},\vec{\lambda})P(\vec{\varphi},\vec{\lambda})\, d^3\varphi\, d^k\lambda}.$$

$P(\vec{\varphi},\vec{\lambda})$ is the *a priori* probability for the cost parameters, $\vec{\varphi}$, and for the parameters of f, $\vec{\lambda}$. As we have no *a priori* information about them, we assume the probabilities to be constant (but see Section 2.4.2). Thus, they come out of the integrals and cancel out, so

$$P(\vec{\varphi},\vec{\lambda}|\vec{x}) = \frac{P(\vec{x}|\vec{\varphi},\vec{\lambda})}{\iint P(\vec{x}|\vec{\varphi},\vec{\lambda})\, d^3\varphi\, d^k\lambda}.$$

$P(\vec{\varphi},\vec{\lambda}|\vec{x})$ is a multidimensional[‡] function that contains all the information we can obtain from the actual positions $\vec{x}$. We will be mostly interested on the information about the cost that we can obtain independently of $\vec{\lambda}$. To do so, we integrate $P(\vec{\varphi},\vec{\lambda}|\vec{x})$ along $\vec{\lambda}$, getting

$$\begin{aligned} P(\vec{\varphi}|\vec{x}) &= \int P(\vec{\varphi},\vec{\lambda}|\vec{x})\, d\vec{\lambda} = \\ &= \frac{\int P(\vec{x}|\vec{\varphi},\vec{\lambda})\, d^k\lambda}{\iint P(\vec{x}|\vec{\varphi},\vec{\lambda})\, d^3\varphi\, d^k\lambda}. \end{aligned}$$

[†]The set of parameters $\vec{\lambda}$ may contain discrete and continuous parameters. In the following we will write multiple integrals along $\vec{\lambda}$, understanding that they are integrals for the continuous parameters and summations for the discrete ones. In the actual implementation of the method we will use three parameters to index function f, so $k = 3$.

[‡]In our case $P(\vec{\varphi},\vec{\lambda}|\vec{x})$ is 6-dimensional: $\vec{\varphi}$ contains the 3 parameters for the cost (ξ, α and β), and $\vec{\lambda}$ contains 3 parameters for function $f_{\vec{\lambda}}$, see Section 2.4.2.

Now we come back to the original notation, substituting $\vec{\varphi}$ for (ξ, α, β):

$$P(\xi, \alpha, \beta|\vec{x}) = \int P(\xi, \alpha, \beta, \vec{\lambda}|\vec{x})\, d^3\lambda = \\ = \frac{\int P(\vec{x}|\xi, \alpha, \beta, \vec{\lambda})\, d^k\lambda}{\iiint d\xi\, d\alpha\, d\beta \int d^k\lambda\, P(\vec{x}|\xi, \alpha, \beta, \vec{\lambda})}.$$

And we substitute Equation (S13), getting the final result:

$$P(\xi, \alpha, \beta|\vec{x}) = \frac{\int f_{\vec{\lambda}}(\vec{x}; \xi, \alpha, \beta)\, d^k\lambda}{\iiint d\xi\, d\alpha\, d\beta \int d^k\lambda\, f_{\vec{\lambda}}(\vec{x}; \xi, \alpha, \beta)}. \tag{S15}$$

2.4.2 Implementation

Actual implementation of the estimator differs from the ideal estimator presented in the previous section in essentially four aspects:

Including a wide range of $f_{\vec{\lambda}}$ functions. It is not possible in practice to take into account all possible forms of the function $f_{\vec{\lambda}}$. We took into consideration a set of functions, trying to make them as general as possible. Any other function can be easily implemented, only increasing the computation time. We implemented six types of functions, and each type depends on two parameters, ν and μ. Therefore, $\vec{\lambda} = (\tau, \nu, \mu)$, where $\tau = 1, 2 \ldots 6$ refers to each of the following functions (names are non-standard).

1. Stretched exponential,

   $$f_{\vec{\lambda}}(W) = \Omega\, e^{-(W/\mu)^\nu}.$$

   Note that exponential and normal probability distributions are particular cases of this, for $\nu = 1$ and $\nu = 2$, respectively.

2. Pseudo-lorentzian type, class 1,

   $$f_{\vec{\lambda}}(W) = \frac{\Omega}{1 + \left(\frac{W}{\mu}\right)^\nu},$$

   of which the lorentzian function is a particular case for $\nu = 2$.

3. Pseudo-lorentzian type, class 2,

$$f_{\vec{\lambda}}(W) = \frac{\Omega}{\left(1 + \left(\frac{W}{\mu}\right)^{\nu}\right)^2}.$$

4. Pseudo-lorentzian type, class 3,

$$f_{\vec{\lambda}}(W) = \frac{\Omega}{\left(1 + \frac{W}{\mu}\right)^{\nu}}.$$

5. Pseudo-lorentzian type, class 4,

$$f_{\vec{\lambda}}(W) = \frac{\Omega}{\left(1 + \left(\frac{W}{\mu}\right)^2\right)^{\nu}},$$

   of which the lorentzian function is a particular case for $\nu = 1$.

6. Power type,

$$f_{\vec{\lambda}} = \Omega \max\left\{1 - \left(\frac{W}{\mu}\right)^{\nu}, 10^{-2}\right\}.$$

In all cases, Ω is a normalization factor, W is the cost, μ is the parameter that controls the spread of the function and ν is a parameter that controls the shape of the function. We assume that μ can take any value between 10^{-11} and 10^{15}, and ν can take any value between 10^{-2} and 10^{4}.[§] The six types of functions are *a priori* assumed to be equiprobable.

Normalization of $P(\vec{x}|\vec{\varphi}, \vec{\lambda})$. In order to normalize $P(\vec{x}|\vec{\varphi}, \vec{\lambda})$, we must evaluate the integral $\int P(\vec{x}|\vec{\varphi}, \vec{\lambda})\, d\vec{x} = \int f_{\vec{\lambda}}(W(\vec{x}; \vec{\varphi}))\, d\vec{x}$. It is not possible in general to compute this integral analytically. To compute it numerically we would need to sample the 279-dimensional space of neuron positions, which is not feasible.

To solve this problem, we neglect the interactions among neurons, assuming that the probability in the direction of one neuron's position does not depend on the

[§]We checked that these extreme values were always regarded as very unlikely at the end of the calculation, suggesting that wider limits are not necessary

position of the other neurons. In practice, this is done by computing the cost in the direction of each neuron, keeping all other neurons fixed in their actual positions, evaluating $f_{\vec{\lambda}}$ for the computed cost, and then normalizing (separately for each neuron). This approximation is especially reasonable in *C. elegans*, because most of the wiring cost is due to connections to muscles and sensors, which do remain fixed in all cases.

Exploration of parameter space. Some of the parameters may span several orders of magnitude (for example, values for α and β may range from 10^{-3} to 10^3). Such wide ranges are better explored in logarithmic scale, which assigns equal weight to each order of magnitude.[¶] Therefore, we chose to estimate the logarithms of α, β, ν and μ, instead of the parameters themselves.[‖] The ranges explored for these parameters were $\alpha \in [10^{-3}, 10^3]$, $\beta \in [10^{-3}, 10^3]$, $\nu \in [10^{-2}, 10^4]$, $\mu \in [10^{-11}, 10^{15}]$. Exponent ξ was explored between 0 and 10.

Numerical evaluation of integrals. Even with the simplifications presented above, it is not feasible to evaluate the integrals accurately over their whole range (note that evaluation of the denominator of equation (S15) requires integration over the 5-dimensional space of ξ, α, β, μ and ν, and that the two latter parameters span several orders of magnitude). Fortunately, it is not necessary to integrate over the whole space, because probability is negligible in the most part of it. Therefore, we proceeded iteratively as follows. We sample the whole space sparsely. Then, we reduce the space neglecting those points where probability is below a certain threshold. This threshold is variable, starting in 10^{-21}, and increasing after each iteration (never above 10^{-4}). After several iterations, we sample the remaining space

[¶]In a linear scale the order of magnitude between 10^2 and 10^3 occupies 90% of the scale, while the three orders of magnitude between 10^{-3} and 10^0 only take about 0.1%. In the logatithmic scale, each of the six orders of magnitude occupies one sixth of the scale.

[‖]Therefore, in practice we assume the *a priori* probability for these parameters to be constant in the logarithmic scale.

more tightly and perform one last calculation that yields the final probability.

2.4.3 Test of the bayesian estimator

We tested the Bayesian estimator on synthetic sets of neuron positions. These sets are generated by stochastic hill climbing with the wiring cost of *C. elegans*, so that we control the values of the cost exponent ξ and the relative weights α and β, and the type of stochastic component added to the cost. Results of these tests are shown in **Figure S6**. In all cases, the Bayesian estimator found the correct values of α and β, **Figure S6B**. The exponent ξ was also estimated correctly, except for a small proportion of the simulations with $\xi = 0.5$, in which the exponent was slightly overestimated, **Figure S6A3**.

The Bayesian estimator finds $\xi = 0.5$ for the experimental data of *C. elegans*. No control test for exponents greater or equal than one resulted in estimated exponent lower than one, **Figure S6A5-8**. This indicates that the estimation of sublinear exponent in the experimental data is robust.

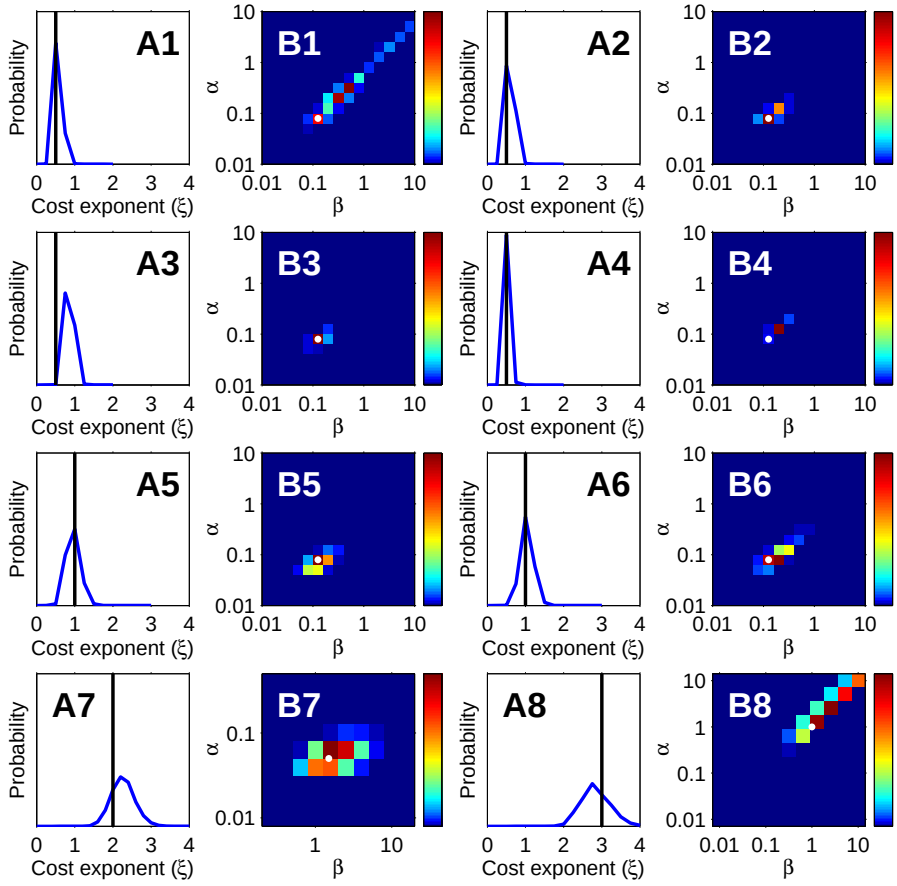


Figure S6: Test of the Bayesian estimator on simulations with known parameters (ξ, α, β). **(A)** Probability yielded by the Bayesian estimator for the cost exponent (ξ). The actual value is marked by the black line. **(B)** Probability found by the Bayesian estimator for α and β. The actual value is marked by the white dot. **Simulations 1 to 4** correspond to $\xi = 0.5$, $\alpha = 0.08$ and $\beta = 0.13$. Simulations 1 to 3 have a normal stochastic component added to the cost. Simulation 4 has uniform stochastic component. **Simulations 5 and 6** have $\xi = 1$, $\alpha = 0.08$, $\beta = 0.13$, and normal stochastic component. **Simulation 7** has $\xi = 2$, $\alpha = 0.05$, $\beta = 1.5$, and normal stochastic component. **Simulation 8** corresponds to $\xi = 3$, $\alpha = \beta = 1$, and normal stochastic component.

2.4.4 Robustness of the wiring cost exponent found by the Bayesian estimator

In this section we will see that the result of $\xi \approx 0.5$ is very robust with respect to changes in α and β. We find that the Bayesian estimator finds a value for the wiring cost exponent near 0.5 for the most part of the values of α and β, including the anatomically-based values, $\alpha = \beta = 1/29.3$, **Figure S7A**. Furthermore, the region of the α-β space where the exponent is around 0.5 matches the region of highest probability, so that higher values of the exponent have probabilities below 10^{-20}, **Figure S7A**.

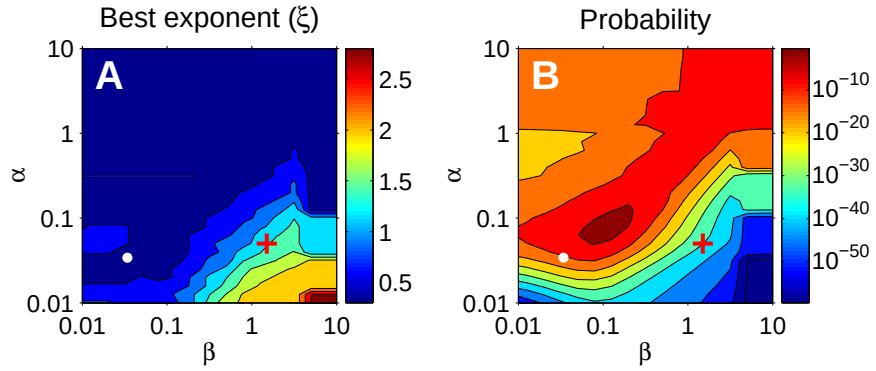


Figure S7: Bayesian estimation of the wiring cost exponent is robust. **(A)** Cost exponent (ξ) most consistent with experimental data for each value of α and β. The exponent is 0.5 for a large region, that includes the most probable values of α, β, the ideal wiring economy values, $\alpha = \beta = 1/29.3$ (white dot), and all the region where probability is higher than 10^{-20}. For the values most consistent with deterministic optimization [3], $\alpha = 0.05$ and $\beta = 1.5$ (red plus sign), the best exponent is 1.3. This is consistent with **Figure 2** of the main text. **(B)** Probability (in logartithmic scale) for α and β.

.

3 Supporting Information for *Escherichia coli*

3.1 Modelling of the metabolic reaction network of *Escherichia coli*

3.1.1 Implementation of the network model

Here we briefly describe the methods used to describe the metabolism of *Escherichia coli*, and find its optimal configuration. A more complete description of these methods can be found in [6, 8, 9, 10]. We focus especially in the particular implementation of these methods using the model presented in **Table S1** (in a separate Excel file).

We can represent all the information about the reactions present in the metabolic network in the stoichiometric matrix **S** [10]. Each row of this matrix represents one metabolite, and each column represents one reaction. The elements of the matrix are the stoichiometric coefficients of the reactions, which are negative when the metabolite is consumed (i.e. it is a reactive) and positive when the metabolite is produced (i.e. it is a product of the reaction). In this study, we use a previously published stoichiometric matrix formed by 98 reactions and 60 metabolites [5]. We removed 24 redundant reactions (included in the original reconstruction because they are catalyzed by different enzymes, but irrelevant for the numerical results of the model). Therefore, 74 reactions remain. This model has the advantage of being relatively small, so that calculations are perfomed fast, while it captures the essential features of the network, providing realistic results. This stoichiometric matrix may be found in table "Matrix" of **Table S1**. A scheme of the reaction network can be found in **Figure S8**.

Typically, each reaction is catalyzed by a different enzyme, so its rate may be regulated independently. We denote ν_i the rate (also called flux) of the i-th reaction. We store all the fluxes in a vector, $\vec{\nu}$. Thus, the product $\mathbf{S} \cdot \vec{\nu}$ gives the net amount of each metabolite that is produced (or consumed, if the figure is negative) per

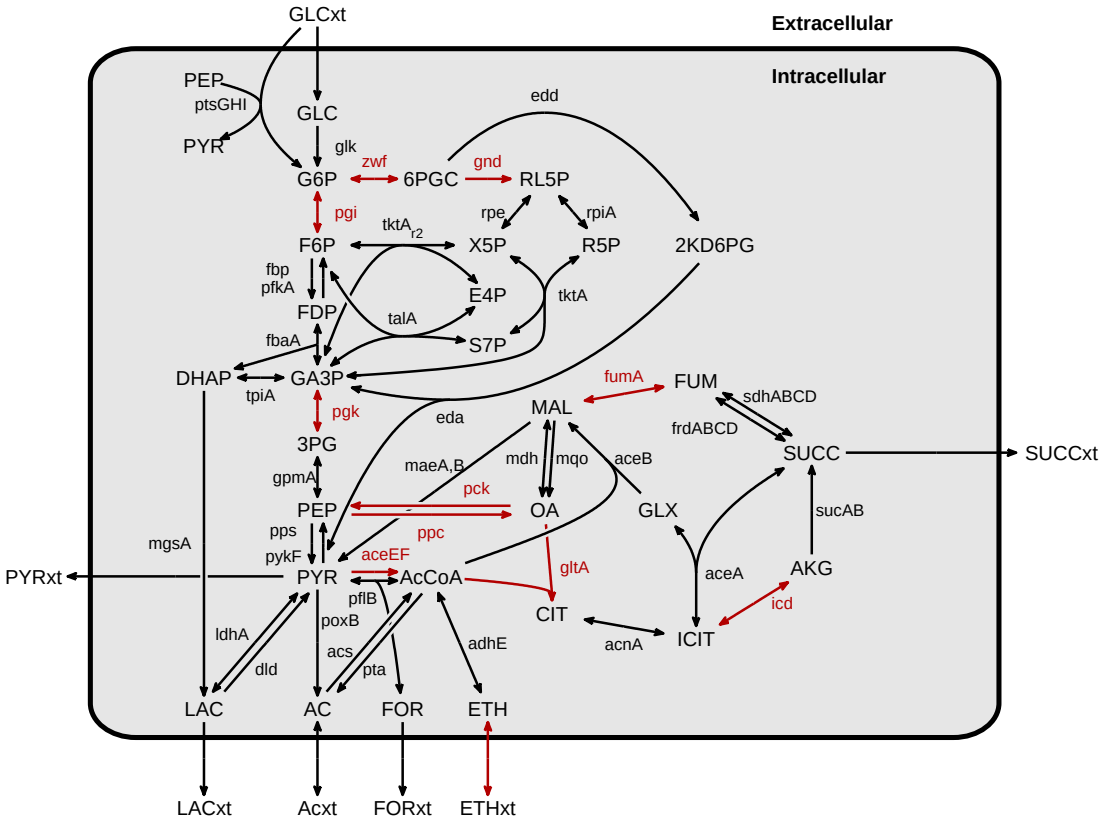


Figure S8: Scheme of the core of the metabolic network, modified from [5]. Reactions whose fluxes are measured experimentally are in red. See **Table S1** (in a separate Excel file) for the meaning of the abbreviations.

unit of time. In the steady state, the net production/consumption of all internal metabolites must be zero, because otherwise metabolites would either accumulate indefinitely in the bacterium's cytoplasm or be created from nothing. Therefore, stoichiometric restrictions and assumption of steady state lead to the fundamental equation

$$\mathbf{S} \cdot \vec{\nu} = \vec{0}. \tag{S16}$$

Equation (S16) is valid for internal metabolites, which cannot be exchanged with the medium. For metabolites that the bacterium can take from the medium (e.g. glucose) the net consumption can be different from zero. Likewise, metabolites that can be secreted to the medium can be produced in excess. This exchange is incorporated to the model as follows [11]. First, for each substance that may be exchanged between the bacterium and the medium, we add an external metabolite to the list (for example, we add "external glucose", GLCxt). External metabolites are metabolites 49-61 in **Table S1**. Then, we add transport reactions to the network. These reactions may be simply the exchange of internal metabolite for external metabolite, or may involve other metabolites (for example, some transport reactions may consume energy in form of ATP, or through other intermediaries). Reactions 62-74 in **Table S1** are transport reactions. Although external metabolites may be produced or consumed in excess, we impose on transport reactions the same condition as we did for internal reactions, Equation (S16). Now, for each external metabolite we add an extra reaction, which is just consumption of the external metabolite and production of nothing. These reactions are denoted *drain reactions* (reactions 75-88 in **Table S1**). These are not proper biochemical reactions, but a convenient way to have external metabolites produced or consumed in excess, while keeping Equation (S16) valid.

Equation (S16), even with the addition of drain reactions, restricts strongly the reaction fluxes that are possible, allowing only those that are consistent with the stoichiometry of all reactions and the condition of steady state. However, we must restrict even further the possible fluxes, because irreversible reactions cannot have

negative fluxes. This is implemented by defining the lower and upper boundary vectors, $\vec{l}$ and $\vec{u}$, respectively. Thus, for reversible reactions $l_i = -\infty$ and $u_i = \infty$, and for irreversible reactions $l_i = 0$ and $u_i = \infty$.** These boundary conditions are also used to fix the value of a given flux, by making $l_i = u_i$ (for example, glucose uptake is fixed in its experimental value, see below).

One major deviation from the stationary state condition is bacterial growth, which is in fact a mayor function of metabolism: Metabolites are constantly accumulating, in order to build new structures for growth and reproduction. This is implemented in the model by adding an extra reaction, which accounts for the production of new biomass [6] (reaction 74 in **Table S1**). This reaction consists of consumption of many metabolites in the proportions observed in bacterial biomass, and production of the virtual metabolite "biomass" (metabolite 60 in **Table S1**), and a small amount of sub-products. Biomass is treated as an external metabolite, with its corresponding drain reaction (reaction 86 in **Table S1**).

Also, the bacterium uses a large amount of energy for purposes different from driving its metabolism. Therefore, the amount of ATP which must be synthesised exceeds the amount needed for the metabolism. This is implemented in the model as *maintenance cost* [6], in two parts: First, growth-associated maintenance (GAM) accounts for the amount of energy which is needed for bacterial growth. It is implemented by increasing the content of ATP of the biomass.†† Second, non-growth-associated maintenance (NGAM) accounts for the energy needed by the bacterium for other processes (such as movement), and which is not proportional to its growth rate. This second kind of maintenance is implemented by adding another virtual external metabolite, the external ATP (metabolite 61 in **Table S1**), with its corresponding drain reaction (reaction 92 in **Table S1**). This drain reaction is

**Some drain reactions are constrained between $-\infty$ and 0. This is because all drain reactions have positive fluxes when the metabolite is secreted, but sometimes the metabolite is actually absorbed from the medium (for example, glucose and oxygen).

††This requires modification of element (2,74) of matrix **S** (see **Table S1**). This is the only modification of the matrix, which is otherwise constant in all calculations.

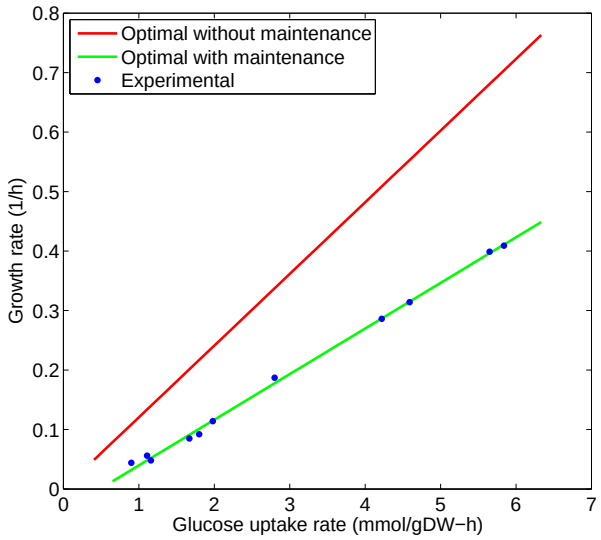


Figure S9: Fit for maintenance requeriment estimation.

constrained to a fixed value, which is the non-growth associated maintenance. The amount of ATP needed for maintenance is calculated by adjusting the values of GAM and NGAM so that growth predicted by the model matches the experimental growth for all experimental conditions (9), see **Figure S9**. From this fit, we obtained NGAM = 12.5 mmolATP/(gDW·h) and GAM = 126 mmolATP/gDW.

We want to find the set of fluxes $\vec{\nu}$ that maximizes the bacterial growth rate, while satisfying the constraints

$$\mathbf{S} \cdot \vec{\nu} = \vec{0}, \qquad \vec{l} \leq \vec{\nu} \leq \vec{u}. \tag{S17}$$

Bacterial growth (or, equivalently, biomass production) is one of the reactions present in matrix **S** (reaction 79 in **Table S1**). Therefore, we want the set of fluxes $\vec{\nu}$ that maximizes one of its elements, while satisfying the constraints (S17). This is a linear optimization problem, which is solved very efficiently by numerical algorithms. We used the open-source GLPK optimization package (GNU Linear Programming Kit, www.gnu.org/software/glpk/). Our routines are based on those of *COBRA* Toolbox [12].

When the objective function is ATP production instead of biomass production,

we change the element of $\vec{\nu}$ to be maximized to the drain reaction for ATP (reaction 87 in **Table S1**). In this case, it makes no difference to include maintenance cost to the model, because the network will in any case produce as much ATP as possible. This choice of objective function has the inconvenient that the optimization diverts all the resources to ATP production, and therefore the predicted biomass production is zero. To avoid this unrealistic result, we fix the biomass production to its experimental value [5].

Different experimental conditions (different growth media, aerobic or anaerobic conditions, etc.) are modelled by adjusting the constraints $(\vec{l}, \vec{u})$ of drain fluxes, controlling the availability of external metabolites [5, 6, 8]. In this case, all experiments correspond to aerobial conditions, so uptake of oxygen was unlimited. We simulated different growth rates by fixing the glucose uptake rate to the experimental value, which is 1.8 mmol/gDW-h for growth rate 0.1 h^{-1}, 2.8 mmol/gDW-h for growth rate 0.2 h^{-1}, 4.4 mmol/gDW-h for growth rate 0.3 h^{-1} and 5.7 mmol/gDW-h for growth rate 0.4 h^{-1} (see bounds for reaction 75 in **Table S1**). For succinate, lactate, pyruvate, formate, acetate, ethanol and NO_2 we allowed unlimited secretion, but no uptake. For NO_3 we allow unlimited uptake but no secretion. For CO_2, we allow unlimited uptake and secretion. See bounds in **Table S1**.

3.1.2 Structure of the deviations from optimality

In this section we describe the calculations that correspond to **Figure 4** of the main text and **Figure S10**.

The first step in testing our theoretical results is to obtain an estimate for the objective function in the direction of each flux. It is not possible to change one flux while keeping all the rest constant, because the linear constraints (Equation S16) would not be satisfied. In order to overcome this problem, we allowed variation of the other fluxes in an interval Δ around their optimal positions. Thus, the procedure to study the objective in the direction of flux i, is the following. First calculate the optimum, $\vec{\nu}^{\text{opt}}$. Then, constrain all fluxes except ν_i to an interval Δ around the

optimum. This means adjusting the boundary vectors $\vec{l}$ and $\vec{u}$, such that

$$l_j = \nu_j^{\text{opt}} - \Delta \quad \text{and} \quad u_j = \nu_j^{\text{opt}} + \Delta, \quad \text{for reversible reactions}$$
$$l_j = \max(\nu_j^{\text{opt}} - \Delta, 0) \quad \text{and} \quad u_j = \nu_j^{\text{opt}} + \Delta, \quad \text{for irreversible reactions,}$$

and for all $j \neq i$, except the objective and those fluxes previously fixed (such as glucose uptake rate, which is fixed in its experimental value). Flux i is constrained to have a value k, by making $l_i = u_i = k$. Then, for each possible value of k, we perform the optimization of the model. When $k = \nu_i^{\text{opt}}$, the result of the optimization will be the same as the global optimum, $\vec{\nu}^{\text{opt}}$, and the objective function will take its maximum value, f^{opt}. For other values of k we obtain the highest value of the objective function compatible with the deviation of ν_i, which in general will be lower than f^{opt}. The interval around the optimal flux for which this calculation is feasible is larger the higher is Δ. Therefore, we must chose Δ at least high enough so that all experimental deviations are included in the feasible intervals. In the limit of $\Delta \to \infty$, this calculation is the same as robustness analysis [10, 12].

The second step is to use Equation (1) of the main text to predict the experimental deviations from the optimum. We expect that the fluxes whose deviations force a fast drop in the objective function will deviate less on average than those fluxes whose deviations are compatible with a high value of the objective. Note that our approximation to the objective in the direction of each flux is essentially the upper bound of the objective. That is, the drop in the objective function when deviating each flux will be as predicted or worse. Therefore, we can be sure that deviations that result in low values of the objective function will indeed have a very low probability.

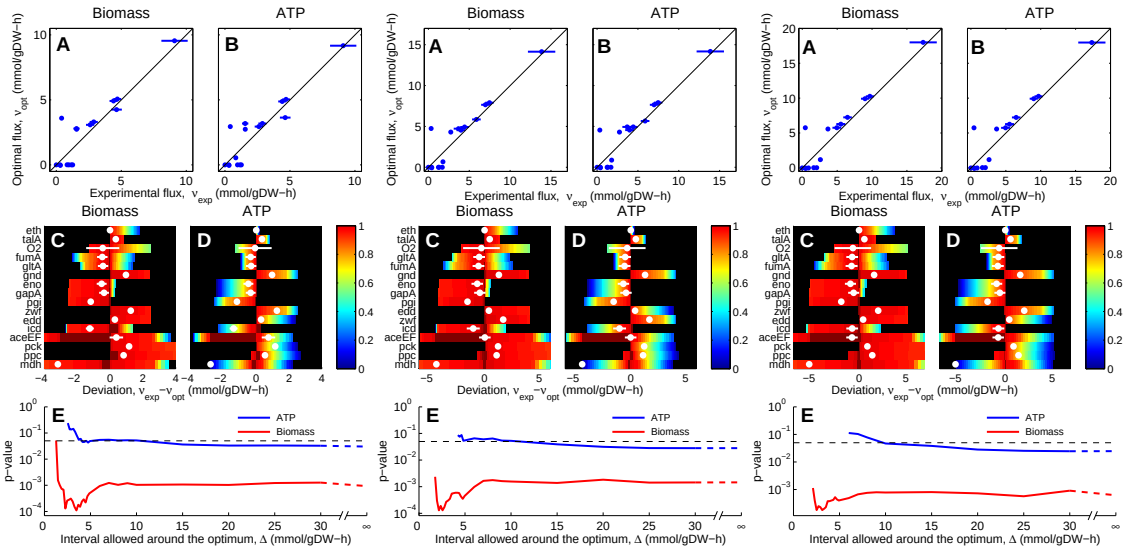
Visual inspection of **Figure 4** of the main text and **Figure S10** reveals that the experimental results match our theoretical results when biomass is the objective function (box C of the Figures), but not for the case when ATP is the objective function (box D of the Figures). In order to see if these results are statistically significant, we developed the following statistical test. First, we compute the upper bound for the objective function in the direction of all fluxes for which experimental

data are available, as described above. Then, we permute the experimental deviations among the fluxes, changing their sign with probability 0.5. In order to prevent that large deviations are assigned to fluxes with small feasible intervals (resulting in non-feasible states), we used the following procedure. We start with the largest deviation, and re-assign it randomly among the fluxes that can accept it. Then, we assign the second-largest among the remaining fluxes that can accept it, and so on. Once all the deviations are re-assigned, we take the value of the estimated objective corresponding to the new deviation for each flux, and average all of them. We compare this average with the one that corresponds to the experimental configuration, and compute the p-value as the proportion of permutations that have a higher average objective than the experimental configuration.

This test confirmed that the experimental results significantly match the theoretical results for the case of growth maximization, and not for the case of ATP maximization. Furthermore, growth maximization works better than ATP maximization for all values of Δ (see box E of **Figure 4** of the main text and **Figure S10**).

3.2 Deviations for other growth rates in *Escherichia coli*

Figure 4 of the main text shows the results for *Escherichia coli* in the case of growth rate 0.1 h^{-1}. We also performed the analysis for other experimental conditions, in which growth rate was 0.2, 0.3 and 0.4 h^{-1}. **Figure S10** shows the results corresponding to these other conditions.



(i) Growth rate 0.2 h^{-1} (ii) Growth rate 0.3 h^{-1} (iii) Growth rate 0.4 h^{-1}

Figure S10: Deviations from optimality in the metabolic network of *E. coli*, for growth rates 0.2, 0.3 and 0.4 h^{-1}. **(A)** Optimal fluxes for maximization of biomass production versus experimental fluxes. Experimental fluxes and their errors are taken from reference [7]. Perfect correspondence of deterministic optimization and data would fall in the diagonal. **(B)** Optimal fluxes for maximization of ATP production versus experimental fluxes. **(C)** Theoretical and experimental deviations from the optimum for biomass production as objective. For each flux, colours show the value of the objective function (relative to the optimum) for each deviation of the flux, with the other fluxes restricted to an interval Δ around their optimum value. $\Delta = 2.54$ for growth rate 0.2 h^{-1}, $\Delta = 4.22$ for growth rate 0.3 h^{-1}, $\Delta = 5.31$ for growth rate 0.4 h^{-1}. Eq. (1) in main text implies that the fluxes should be at the red regions. White dots are located at the experimental deviations. **(D)** Theoretical and experimental deviations for ATP production as objective, same Δ as in box (C). **(E)** p-value from significance analysis, for all possible values of Δ. Red line: Significance of the theoretical results using biomass production as objective. Blue line: Using ATP production as objective. Black dashed line: $p = 0.05$ significance line.

References

[1] Chen, B.L., Hall, D.H., Chklovskii, D.B. Wiring optimisation can relate neuronal structure and function. *Proc. Natl. Acad. Sci. USA* 103, 4723-4728 (2006) (also see its Supplementary Material)

[2] www.wormatlas.org

[3] Pérez-Escudero, A., de Polavieja, G. Optimally wired subnetwork determines neuronanatomy of Caenorhabditis elegans, *Proc. Natl. Acad. Sci. USA* 104, 17180-17185 (2007)

[4] Chklovskii, D.B. Exact Solution for the Optimal Neuronal Layout Problem. *Neural Computation* 16, 2067-2078 (2004)

[5] Schuetz, R., Kuepfer, L., Sauer, U. Systematic evaluation of objective functions for predicting intracellular fluxes in Escherichia coli. *Molecular Systems Biology* 3:119 (2007)

[6] Varma, A. & Palsson, B.O. Stoichiometric Flux Balance Models Quantitatively Predict Growth and Metabolic By-Product Secretion in Wild-Type Escherichia coli W3110. *Applied and Environmental Microbiology*, 60(10):3724-3731 (1994)

[7] Nanchen, A., Schicker, A. & Sauer, U. Nonlinear Dependency of Intracellular Fluxes on Growth Rate in Miniaturized Continuous Cultures of Escherichia coli. *Applied and Environmental Microbiology* 72(2):1164-1172 (2006)

[8] Varma, A. & Palsson, B.O. Metabolic capabilities of *Escherichia coli*: I. Synthesis of Biosynthetic Precursors and Cofactors. *J Theor Biol* 165, 477-502 (1993)

[9] Varma, A. & Palsson, B.O. Metabolic capabilities of *Escherichia coli*: II. Optimal growth patterns. *J Theor Biol* 165, 503-522 (1993)

[10] Palsson, B.O. *Systems Biology: Properties of reconstructed networks*. Cambridge University Press (2006)

[11] Feist A.M., Henry C.S., Reed J.L., Krummenacker M.K., Joyce A.R., Karp P.D., Broadbelt L.J., Hatzimanikatis V. & Palsson B.O. A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Molecular Systems Biology* 3:121 (2007)

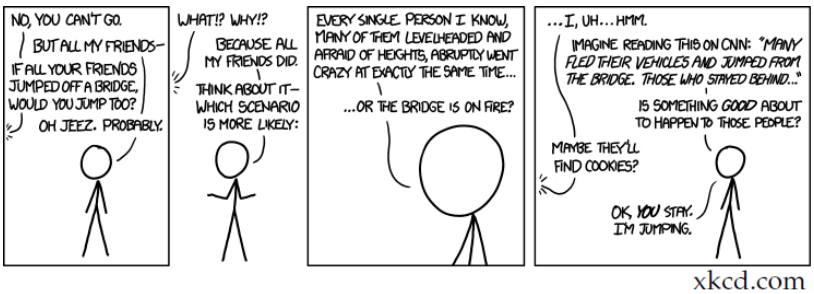
[12] Becker, S.A., Feist, A.M., Mo, M.L, Hannum, G., Palsson, B.O. & Herrgard, M.J. (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA toolbox. *Nature Protocols* 2(3):727-738

4
A Bayesian model for collective decision-making

La gente no es tonta.

(People are not stupid)

Adela Escudero Berián



OPEN ACCESS Freely available online

PLoS COMPUTATIONAL BIOLOGY

Collective Animal Behavior from Bayesian Estimation and Probability Matching

Alfonso Pérez-Escudero[1,2]*, Gonzalo G. de Polavieja[1,2]*

1 Instituto Cajal, Consejo Superior de Investigaciones Científicas, Madrid, Spain, **2** Department of Theoretical Physics and Instituto "Nicolás Cabrera" de Física de Materiales, Universidad Autónoma de Madrid, Madrid, Spain

Abstract

Animals living in groups make movement decisions that depend, among other factors, on social interactions with other group members. Our present understanding of social rules in animal collectives is mainly based on empirical fits to observations, with less emphasis in obtaining first-principles approaches that allow their derivation. Here we show that patterns of collective decisions can be derived from the basic ability of animals to make probabilistic estimations in the presence of uncertainty. We build a decision-making model with two stages: Bayesian estimation and probabilistic matching. In the first stage, each animal makes a Bayesian estimation of which behavior is best to perform taking into account personal information about the environment and social information collected by observing the behaviors of other animals. In the probability matching stage, each animal chooses a behavior with a probability equal to the Bayesian-estimated probability that this behavior is the most appropriate one. This model derives very simple rules of interaction in animal collectives that depend only on two types of reliability parameters, one that each animal assigns to the other animals and another given by the quality of the non-social information. We test our model by obtaining theoretically a rich set of observed collective patterns of decisions in three-spined sticklebacks, *Gasterosteus aculeatus*, a shoaling fish species. The quantitative link shown between probabilistic estimation and collective rules of behavior allows a better contact with other fields such as foraging, mate selection, neurobiology and psychology, and gives predictions for experiments directly testing the relationship between estimation and collective behavior.

Citation: Pérez-Escudero A, de Polavieja GG (2011) Collective Animal Behavior from Bayesian Estimation and Probability Matching. PLoS Comput Biol 7(11): e1002282. doi:10.1371/journal.pcbi.1002282

Editor: Iain D. Couzin, Princeton University, United States of America

Received April 7, 2011; **Accepted** October 5, 2011; **Published** November 17, 2011

Funding: This work was funded by MICIIN (Spain) as Plan Nacional (http://www.micinn.es) and as partners of the ERASysBio+ initiative supported under the EU ERA-NET Plus scheme in FP7 (http://www.erasysbio.net/), and by Biociencia program (CAM, Spain) (http://www.madrimasd.org/). A.P-E. acknowledges a FPU fellowship from MICINN (Spain). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: alfonso.perez.escudero@cajal.csic.es (APE); gonzalo.polavieja@cajal.csic.es (GGdP)

Introduction

Animals need to make decisions without certainty in which option is best. This uncertainty is due to the ambiguity of sensory data but also to limited processing capabilities, and is an intrinsic and general property of the representation that animals can build about the world. A general way to make decisions in uncertain situations is to make probabilistic estimations [1,2]. There is evidence that animals use probabilistic estimations, for example in the early stages of sensory perception [3–11], sensory-motor transformations [12–14], learning [15–17] and behaviors in an ecological context such as strategies for food patch exploitation [18–20] and mate selection [21], among others [13,17,21,22].

An additional source of information about the environment may come from the behavior of other animals (social information) [23–28]. This information can have different degrees of ambiguity. In particular cases, the behavior of conspecifics directly reveals environmental characteristics (for example, food encountered by another individual informs about the quality of a food patch). Cases in which social information correlates well with the environmental characteristic of interest have been very well studied [29–37]. But in most cases social information is ambiguous and potentially misleading [26,38]. In spite of this ambiguity, there is evidence that in some cases such as predator avoidance [39,40] and mate choice [41], animals use this kind of information.

Social animals have a continuous flow of information about the environment coming from the behaviours of other animals. It is therefore possible that social animals use it at all times, making probabilistic estimations to counteract its ambiguity. If this is the case, estimation of the environment using both non-social and social information might be a major determinant of the structure of animal collectives. In order to test this hypothesis, we have developed a Bayesian decision-making model that includes both personal and social information, that naturally weights them according to their reliability in order to get a better estimate of the environment. All members of the group can then use these improved estimations to make better decisions, and collective patterns of decisions then emerge from these individuals interacting through their perceptual systems.

We show that this model derives social rules that economically explain detailed experiments of decision-making in animal groups [42,43]. This approach should complement the empirical approach used in the study of animal groups [42–47], finding which mathematical functions should correspond to each experimental problem and to propose experiments relating estimation and collective motion. The Bayesian structure of our model also builds a bridge between the field of collective behavior and other fields of animal behavior, such as optimal foraging theory [18–22] and others [21,22]. Further, it explicitly includes in

Author Summary

Animals need to act on uncertain data and with limited cognitive abilities to survive. It is well known that our sensory and sensorimotor processing uses probabilistic estimation as a means to counteract these limitations. Indeed, the way animals learn, forage or select mates is well explained by probabilistic estimation. Social animals have an interesting new opportunity since the behavior of other members of the group provides a continuous flow of indirect information about the environment. This information can be used to improve their estimations of environmental factors. Here we show that this simple idea can derive basic interaction rules that animals use for decisions in social contexts. In particular, we show that the patterns of choice of *Gasterosteus aculeatus* correspond very well to probabilistic estimation using the social information. The link found between estimation and collective behavior should help to design experiments of collective behavior testing for the importance of estimation as a basic property of how brains work.

a natural way different cognitive abilities, making more direct contact with neurobiology and psychology [3–10,17].

Results

Estimation model

We derived a model in which each individual decides from an estimation of which behavior is best to perform. These behaviors can be to go to one of several different places, to choose among some behaviors like forage, explore or run away, or any other set of options. For clarity, here we particularize to the case of choosing the best of two spatial locations, x and y (see *Text S1* for more than two options). 'Best' may correspond to the safest, the one with highest food density or most interesting for any other reasons. We assume that each decision maker uses in the estimation of the best location both non-social and social information. Non-social information may include sensory information about the environment (i.e. shelter properties, potential predators, food items), memory of previous experiences and internal states. Social information consists of the behaviors performed by other decision-makers. Each individual estimates the probability that each location, say y, is the best one, using its non-social information (C) and the behavior of the other individuals (B),

$$P(Y|C,B), \tag{1}$$

where Y stands for 'y is the best location'. $P(X|C,B)=1-P(Y|C,B)$, because there are only two locations to choose from. We can compute the probability in Eq. **1** using Bayes' theorem,

$$P(Y|C,B)=\frac{P(B|Y,C)P(Y|C)}{P(B|X,C)P(X|C)+P(B|Y,C)P(Y|C)}. \tag{2}$$

By simply dividing numerator and denominator by the numerator we find an interesting structure,

$$P(Y|C,B)=\frac{1}{1+a\,S}, \tag{3}$$

where

$$a=\frac{P(X|C)}{P(Y|C)} \tag{4}$$

and

$$S=\frac{P(B|X,C)}{P(B|Y,C)}. \tag{5}$$

Note that a does not contain any social information so it can be understood as the "non-social term" of the estimation. We can also understand S as the "social term" because it contains all the social information, although is also depends on the non-social information C. The non-social term a is the likelihood ratio for the two options given only the non-social information. This kind of likelihood ratio is the basis of Bayesian decision-making in the absence of social information [5,11–14]. Eq. **3** now tells us that this well known term interacts with the social term S simply through multiplication.

We are seeking a model based on probabilistic estimation that can simultaneously give us insight into social decision-making and fit experimental data. For this reason we simplify the model by assuming that the focal individual does not make use of the correlations among the behaviour of others, but instead assumes their behaviours to be independent of each other. This is a strong hypothesis but allows us to derive simple explicit expressions with important insights. The section 'Model including dependencies' at the end of Results shows that this assumption gives a very good approximation to a more complete model that takes into account these correlations.

The assumption of independence translates in that the probability of a given set of behaviors is just the product of the probabilities of the individual behaviors. We apply it to the probabilities needed to compute S in Eq. **5**, getting

$$P(B|Y,C)=Z\prod_{i=1}^{N}P(b_i|Y,C), \tag{6}$$

where B is the set of all the behaviors of the other N animals at the time the focal individual chooses, $B=\{b_i\}_{i=1}^{N}$, and b_i denotes the behavior of one of them, individual i. Z is a combinatorial term counting the number of possible decision sequences that lead to the set of behaviors B, that will cancel out in the next step. Substituting Eq. **6** and the corresponding expression for $P(B|X,C)$ into Eq. **5**, we get

$$S=\prod_{i=1}^{N}\frac{P(b_i|X,C)}{P(b_i|Y,C)}. \tag{7}$$

Instead of an expression in terms of as many behaviors as individuals, it may be more useful to consider a discrete set of behavioral classes. For example, in our two-choice example, these behavioral classes may be 'go to x' (denoted β_x), 'go to y' (β_y) and 'remain undecided' (β_u). Frequently, these behavioral classes (or simply 'behaviors') will be directly related to the choices, so that each behavior will consist of choosing one option. For example, behaviors β_x and β_y are directly related to choices x and y, respectively. But there may be behaviors not related to any option as the case of indecision, β_u, or related to choices in an indirect way. These behaviors can still be informative because they may be more consistent with one of the options being better than the other (for example, indecision may increase when there is a predator, so

the presence of undecided individuals may bias the decision against the place where the non-social information suggests the presence of a predator). Let us consider L different behavioral classes, $\{\beta_k\}_{k=1}^{L}$. We do not here consider individual differences for animals performing the same behavior (say, behavior β_1), so they have the same probabilities $P(\beta_1|X,C)$ and $P(\beta_1|Y,C)$. Thus, if for example the n_1 first individuals are performing behavior β_1, we have that $\Pi_{i=1}^{n_1}\frac{P(b_i|X,C)}{P(b_i|Y,C)}=\left(\frac{P(\beta_1|X,C)}{P(\beta_1|Y,C)}\right)^{n_1}$. We can then write Eq. **7** as

$$S=\prod_{k=1}^{L}s_k^{n_k}, \tag{8}$$

where n_k is the number of individuals performing behavior β_k, and

$$s_k=\frac{P(\beta_k|X,C)}{P(\beta_k|Y,C)}. \tag{9}$$

The term s_k is the probability that an individual performs behavior β_k when x is the best option, over the probability that it performs the same behavior when y is the best choice. The higher s_k the more reliably behavior β_k indicates that x is better than y, so we can understand s_k as the reliability parameter of behavior β_k. If $s_k=\infty$, observing behavior β_k indicates with complete certainty that x is the best option, while for $s_k=1$ behavior β_k gives no information. For $s_k<1$, observing behavior β_k favors y as the best option, and more so the closer it is to 0. Note that $P(\beta_k|X,C)$ and $P(\beta_k|Y,C)$ are not the actual probabilities of performing behavior β_k, but estimates of these probabilities that the deciding animal uses to assess the reliability of the other decision-makers. These estimates may be 'hard-wired' as a result of evolutionary adaptation, but may also be subject to change due to learning.

To summarize, using Eqs. **3** and **8**, the probability that y is the best choice, given both social and non-social information is

$$P(Y|C,B)=\left(1+a\prod_{k=1}^{L}s_k^{n_k}\right)^{-1}, \tag{10}$$

with a in Eq. **4** and s_k in Eq. **9**.

Decision rule: Probability matching

We have so far only considered the perceptual stage of decision-making, in which the deciding individual estimates the probability that each behavior is the best one. Now it must decide according to this estimation. A simple decision rule would be to go to y when $P(Y|C,B)$ is above a certain threshold. This rule maximizes the amount of correct choices when the probabilities do not change [48], but is not consistent with the experimental data considered in this paper. Applying this deterministic rule strictly, without any noise sources, one would obtain that all individuals behave exactly in the same way when facing the same stimuli, but in the experiments considered here this is not the case. Instead, we used a different decision rule called probability matching, that has been experimentally observed in many species, from insects to humans [49–55]. According to this rule an individual chooses each option with a probability that is equal to the probability that it is the best choice. Therefore, in our case the probability of going to y (P_y), is the same as the estimated probability that y is the best location ($P(Y|C,B)$), so

$$P_y=P(Y|C,B). \tag{11}$$

Probability matching does not maximize the amount of right choices if we assume that the probabilities stay always the same, but in many circumstances it can be the optimal behavior, such as when there is competition for resources [56,57], when the estimated probabilities are expected to change due to learning [53,55], or for other reasons [53,58].

Finally, using Eqs. 10 and 11 we have that the probability that the deciding individual goes to y is

$$P_y=\left(1+a\prod_{k=1}^{L}s_k^{n_k}\right)^{-1}. \tag{12}$$

The assumption of probability matching has the advantage that the final expression for the decision in Eq. **12** is identical to the one given by Bayesian estimation in Eq. **10**, with no extra parameters. Alternative decision rules could be noisy versions of the threshold rule, but at the price of adding at least one extra parameter to describe the noise. Also, decision rules might not depend on estimation alone, but also on other factors or constraints. These more complicated rules fall beyond the scope of this paper.

In the following sections, we particularize Eq. **12** to different experimental settings to test its results against existing rich experimental data sets that have previously been fitted to different mathematical expressions [42,43].

Symmetric set-up

We first considered the simple case of two identical equidistant sites, x and y, Fig. 1*A*. For a set-up made symmetric by experimental design there is no true best option. But deciding individuals must act, like for any other case, using only their incomplete sensory data to make the best possible decision. Even when non-social sensory data indicates no relevant difference between the two sites, the social information can bias the estimation of the best option to one of the two sites.

Using Eq. **12** and that the three possible behaviors are 'go to x' (β_x), 'go to y' (β_y) and 'remain undecided' (β_u), we obtain

$$P_y=\left(1+a\,s_x^{n_x}s_y^{n_y}s_u^{N-n_x-n_y}\right)^{-1}, \tag{13}$$

where n_x and n_y are the number of individuals that have already chosen x and y, respectively, and $N+1$ is the size of the group containing our focal individual and other N animals. As the set-up is symmetric, the sensory information available to the deciding individual is the same for both options so $P(X|C)=P(Y|C)$ and then $a=1$ according to Eq. **4**. Also, since indecision is not related to any particular choice, symmetry imposes $P(\beta_u|X,C)=P(\beta_u|Y,C)$, so indecision is not informative, $s_u=1$ (Eq. **9**). For the other two behaviors, going to x (β_x) and going to y (β_y), Eq. **9** gives

$$\begin{aligned}s_x&=\frac{P(\beta_x|X,C)}{P(\beta_x|Y,C)}\\ s_y&=\frac{P(\beta_y|X,C)}{P(\beta_y|Y,C)}.\end{aligned} \tag{14}$$

$P(\beta_x|X,C)$ and $P(\beta_y|Y,C)$ are the estimated probabilities of making the right choice, that is, going to x when x is the best option, or going to y when y is the best option. Since in this case the sensory information is identical for both options, the probability of making the correct choice must be the same for both options, $P(\beta_x|X,C)=P(\beta_y|Y,C)$. An analogous argument

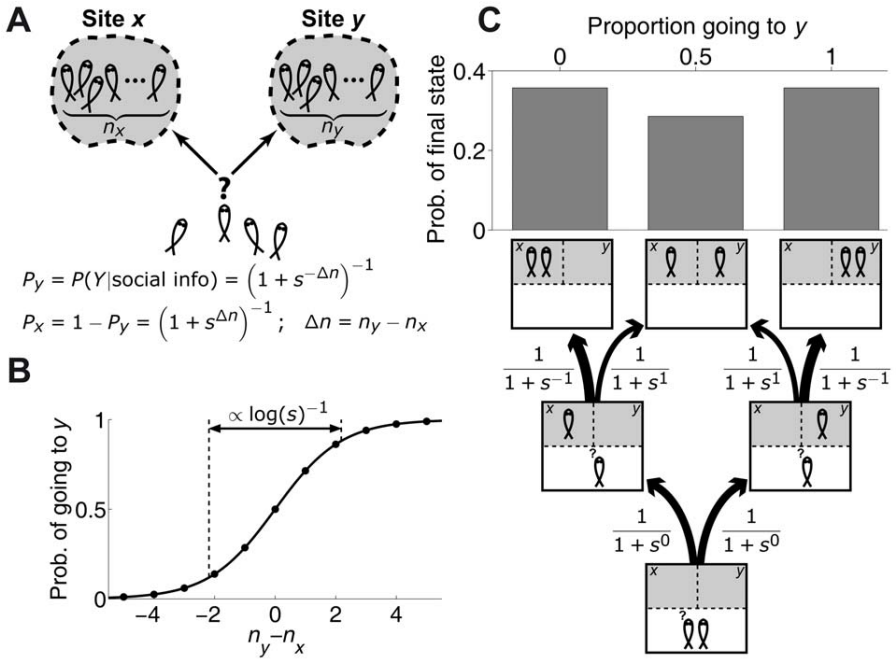


Figure 1. Model with individuals estimating which of two identical places is best. (*A*) Schematic diagram of individuals choosing between two identical locations x and y when there are already n_x (n_y) individuals at x (y). (*B*) Probability of going to y as a function of the difference between the number of individuals at y and x, Eq. **17**. (*C*) Sequential application of the behavioural rule in Eq. **17** with $s=2.5$, for the simple case of a group of two individuals (bottom). The width of the arrows is proportional to the probability of each transition. The 3 possible final configurations, with different proportion of individuals going to y (0, 0.5 and 1), have different probabilities of taking place, with both fish together at x or y being more probable than a group split (top).
doi:10.1371/journal.pcbi.1002282.g001

holds for the incorrect choices, $P(\beta_x|Y,C)=P(\beta_y|X,C)$, giving

$$s_x = 1/s_y. \tag{15}$$

In cases in which $s_x=1/s_y$, we find it convenient to express reliability more generally as

$$s \equiv s_x = 1/s_y, \tag{16}$$

which is the ratio of the probability of making the correct choice and the probability of making a mistake, for both behaviors. Using this definition and given that $a=a_u=1$, Eq. **13** reduces to

$$P_y = \left(1+s^{-\Delta n}\right)^{-1}, \tag{17}$$

with the variable $\Delta n \equiv n_y - n_x$. Eq. **17** describes a sigmoidal function that is steeper the higher the higher the value of s (Fig. 1*B*). Therefore, for very reliable behaviors (high s, meaning individuals that are much more likely to make correct choices than erroneous ones), P_y grows fast with Δn and the deciding individual then goes to y with high probability when taking into account the behaviors of only very few individuals.

The behavior of the group is obtained by applying the decision rule in Eq. **17** sequentially to each individual (see *Methods*). After each behavioural choice, we update the number of individuals at x and y, using the new n_x and n_y for the next deciding individual (Fig. 1*C*, bottom). Repeating this procedure for all the individuals in the group, we can compute the probability for each possible final outcome of the experiment (Fig. 1*C*, top).

The relevance of the symmetric case is that the model has a single parameter and a single variable, enabling a powerful comparison against experimental data. We tested the model using an existing rich data set of collective decisions in three-spined sticklebacks [42], a shoaling fish species. This data set was obtained using a group of N_{tot} fish choosing between two identical refugia, one on their left and another one on their right (Fig. 2*A*), equivalent to locations x and y in the model (Fig. 1*A*). At the start of the experiment, m_x (m_y) replica fish made of resin were moved along lines on the left (right) towards the refugia (Fig. 2*A*). The experimental results consisted on the statistics of collective decisions between the two refugia for 19 different cases using different group sizes $N_{\text{tot}}=2$, 4 or 8 and different numbers of replicas going left and right, $m_x : m_y = \{1{:}1, 2{:}2, 0{:}1, 1{:}2, 0{:}2, 1{:}3, 0{:}3\}$ (Fig. 2*B*, blue histograms). To compare against these experimental data, we calculated the probability of finding a collective pattern applying the individual behavioural rule in Eq. **17** iteratively over each fish for the 19 experimental settings. We found a good fit of the model to the experimental data using for the 19 graphs the same value $s=2.2$ (Fig. 2*B*, red line). The model is robust, with good fits in the interval $s=2-4$ (Fig. 3, red line).

Despite the simplicity of the behavioral rule in Eq. **17**, it reproduces the experimental results, including the dependence on the total number of fish N_{tot}, even though the rule is independent of this parameter, except for determining the range of possible values of Δn. The dependence of the final distributions on N_{tot} emerges from the application of the rule to the N_{tot} individuals in the group, as is illustrated in Fig. 4. Each small box represents a state of the system in which $n_x : n_y$ fish have already decided to go to x and y, respectively. The lines connecting each box with another two boxes on top represent the decision made by the next

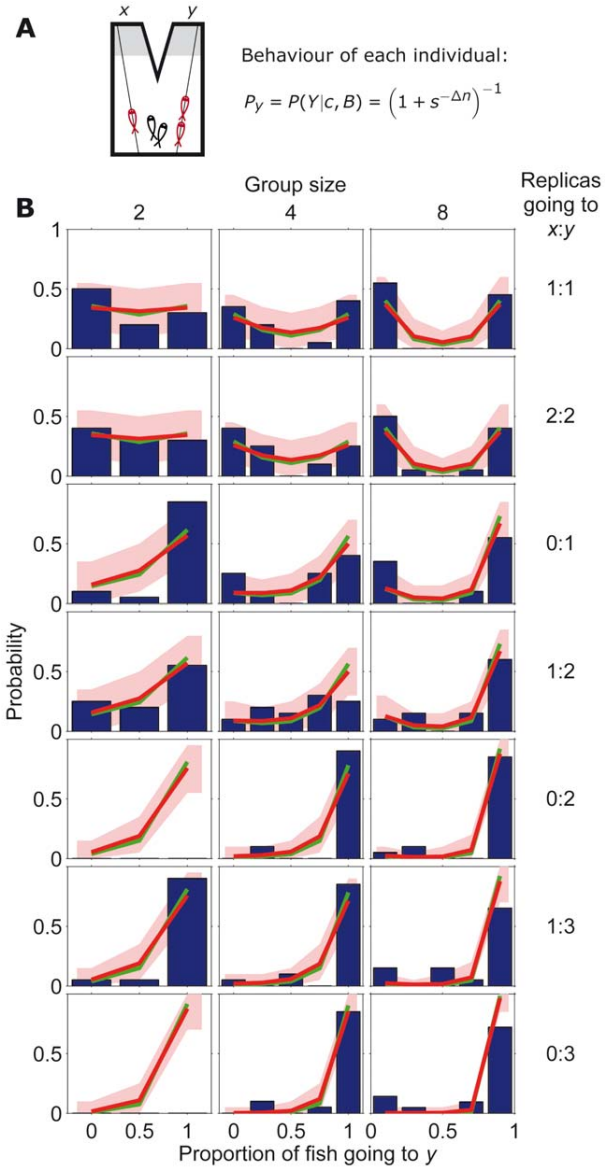


Figure 2. Comparison between model and stickleback choices in symmetric set-up. (*A*) Schematic diagram of symmetric set-up with a group of sticklebacks (in black) choosing between two identical refugia and with different numbers of replica fish (in red) going to x and y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from the model in Eq. **17** in the main text (red line using reliability parameter $s=2.2$; red region: 95% confidence interval; green line with $s=2.5$). Different graphs correspond to different stickleback group sizes and different number of replicas going to x and y.
doi:10.1371/journal.pcbi.1002282.g002

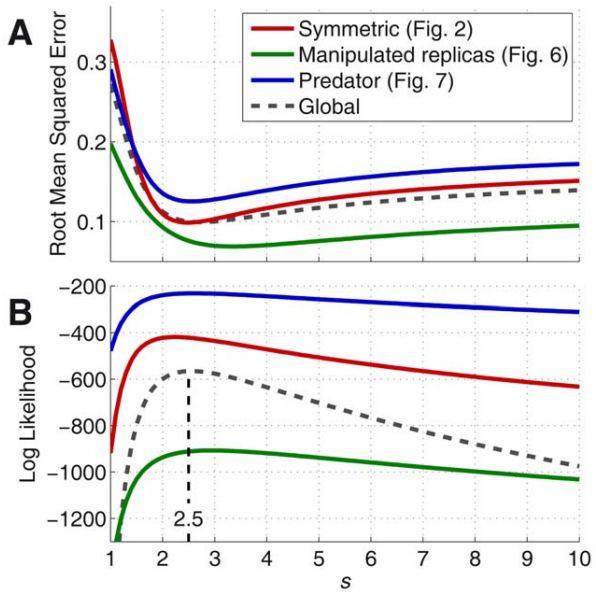


Figure 3. Goodness of fit for different values of the reliability (s). Red: Symmetric case (plots in Fig. 2). **Green:** Case with different replicas at each side (plots in Fig. 6. The ratios s_r/s_R are re-optimized for each value of s). **Blue:** Asymmetric set-up with predator on one side (plots in Fig. 7 ; Parameter a is re-optimized for each value of s). (*A*) Root mean squared error between the data and the probabilities predicted by the model. Grey dashed line shows the mean RMSE for the three cases. The absolute values for each case depend on the shape of the data and are not comparable, only the trends and the position of the minima should be compared. (*B*) Logarithm of the probability that the data come from the model. The height of each curve depends on the number of data for each experiment, only the trend and the position of the maxima should be compared. Grey dashed line shows the sum of the three coloured lines, but shifted by 1000 so that it fits on the scale. The peak of this global probability indicates the value of s that best fits the three datasets ($s=2.5$).
doi:10.1371/journal.pcbi.1002282.g003

deciding individual, that takes the system to the next state. The width of the lines is proportional to the probability of the decision. As more individuals decide, the central states become less likely simply because they accumulate more unlikely decisions. Therefore, the U-shape or J-shape becomes more pronounced for larger groups, even though the individual decision rule in Eq. **17** is independent of the total number of individuals N_{tot}.

Group decision-making in three-spined sticklebacks shows a single type of distribution in which probability is minimum at the center and increases monotonically towards the edges, denoted here as U-shaped distribution (or J-shaped when there is a bias to one of the two options). However, the model in Eq. **17** also gives two other types of distributions, Fig. 5*A*. For non-social behavior ($s=1$) the histogram is bell-shaped due to combinatorial effects. However, a bell-shape is also compatible with social animals for a certain range of s and group size (white region on the bottom-left of Fig. 5*A*). For higher values of s, the histograms are M-shaped, with two maxima located between the center and the sides (region coloured in black and blue in Fig. 5*A*). However, the M shape becomes clear only with enough number of bins because the drop in probability near the edge or at the center of the distribution disappears when binning is too coarse, producing a bell-shaped or U-shaped histogram, Fig. 5*B*. This is an important practical issue, because the amount of data that can be collected rarely allows for more than 5 bins. The colorscale in Fig. 5*A* reflects the number of bins needed to observe the M shape (black has been reserved for exactly 5 bins). For high values of s, the histograms are U-shaped (white region on the top of Fig. 5*A*). Also, all the M-region above the black zone becomes of type U when the binning is too coarse.

An interesting prediction of our model is that, for a given number of bins, the shape of the distribution of choices changes with the number of decided individuals, and the dynamics of this change depends on s. For high values of s, the probability is U-

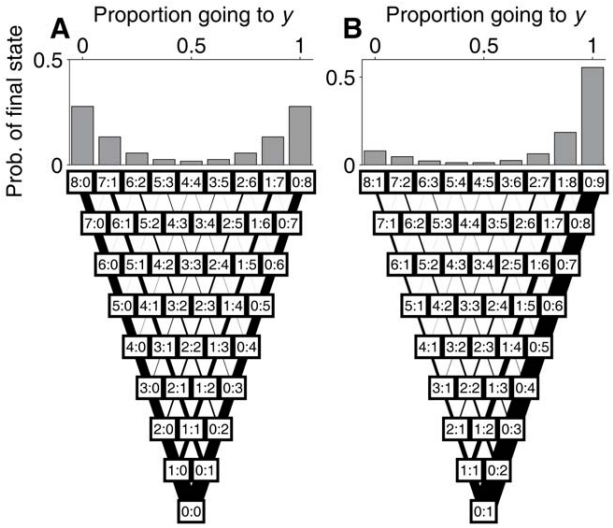


Figure 4. Illustration of the decision-making process in the model. Bottom: Decision-making process according to Eq. **17** (with $s=2.5$). Time runs from bottom to top. Each box represents a state with a given number of fish having already decided x or y ($n_x : n_y$). Each state can lead to another two states in the following time step, depending on whether the focal fish decides to go to x or y. The width of the lines connecting states is proportional to the probability of that transition (equal to the probability of the prior state times the probability of the focal fish making the decision that leads to the later one). **Top:** Probability of each state after 8 fish have made their decisions. (*A*) Case with no replicas, in which the final outcome is U-shaped. (*B*) Case with one replica going to y (so initial state is already 0:1), in which the final outcome is J-shaped.
doi:10.1371/journal.pcbi.1002282.g004

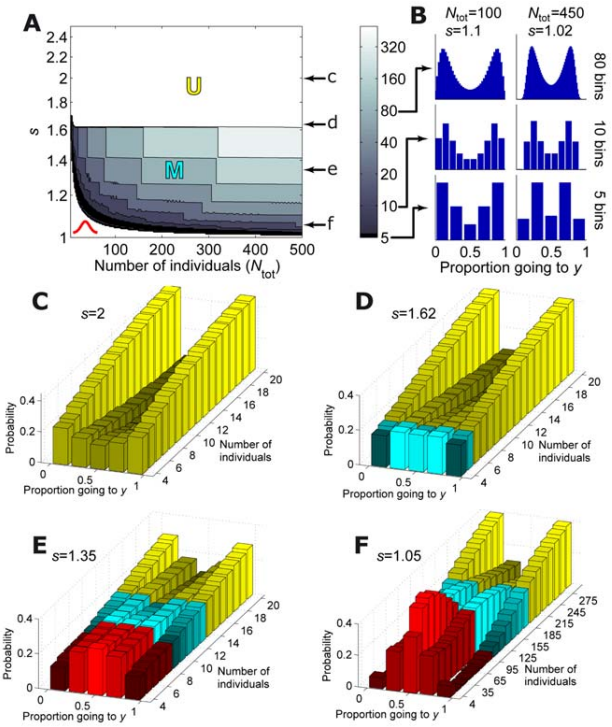


Figure 5. Types of distributions and dynamics for different values of reliability parameter s and group size. (*A*) Shape of histogram of final configurations as a function of s and the group size. Bell-shaped: white region on the bottom-left. M-shaped: region coloured in black and blue. As the observation of the M shape depends on the number of bins, the colorscale reflects the number of bins needed to observe the M shape (black has been reserved for exactly 5 bins). U-shape: white region on the top. Also, all the M-region above the black zone becomes U when the binning is too coarse. There is also a small region below the black zone where the M shape becomes a bell shape when the binning is too coarse. (*B*) Dependence of the apparent shape on the number of bins: Top, 80 bins. Middle, 10 bins. Bottom, 5 bins. On the left, a probability that seems U-shaped for 5 bins, but is M shaped for a higher number of bins. On the right, a probability that stays M-shaped for any number of bins. (*C–F*) Dynamics of the probability as the number of individuals increases for (*C*) $s=2$, (*D*) $s=1.62$, (*E*) $s=1.35$ and (*F*) $s=1.05$.
doi:10.1371/journal.pcbi.1002282.g005

shaped from the beginning and becomes steeper as more individuals decide (as is the case for the stickleback dataset), Fig. 5*C*. For lower values of s, we observe M-shaped distributions for the first individuals and then U-shaped ones when more individuals decide, Fig. 5*D*. For even lower values of s, we observe bell-shaped distributions for the first individuals, then M-shaped and finally U-shaped, Fig. 5*E,F*.

Symmetric set-up with modified replicas of animals

An interesting modification of the experimental set-up consists in using replicas of the animals that we can modify to potentially alter their reliability estimated by the animals. We considered the particular case, motivated by experiments in [43], of two types of modified replicas with different characteristics (for example, fat or thin), Fig. 6*A*. We considered 7 behaviors: 'animal goes to x' (β_{fx}), 'animal goes to y' (β_{fy}), 'most attractive replica goes to x' (β_{Rx}), 'most attractive replica goes to y' (β_{Ry}) 'least attractive replica goes to x' (β_{rx}), 'least attractive replica goes to y' (β_{ry}), and 'animal remains undecided' (β_{fu}). The probability of going to y in Eq. **12** then reduces to

$$P_y=\left(1+a\,s_{fx}^{n_{fx}}s_{fy}^{n_{fy}}s_{Rx}^{n_{Rx}}s_{Ry}^{n_{Ry}}s_{rx}^{n_{rx}}s_{ry}^{n_{ry}}s_{fu}^{N_f-n_{fx}-n_{fy}}\right)^{-1}, \quad (18)$$

where subindex 'f' refers to real fish and 'R' ('r') to replicas of the most (least) attractive type. As in the previous section, symmetry imposes that $a=1$ and $s_{fu}=1$. It also imposes the following relations between the reliability parameters, $s_f\equiv s_{fx}=1/s_{fy}$, $s_R\equiv s_{Rx}=1/s_{Ry}$, $s_r\equiv s_{rx}=1/s_{ry}$. Therefore,

$$P_y=\left(1+s_f^{-\Delta n_f}s_R^{-\Delta n_R}s_r^{-\Delta n_r}\right)^{-1}, \quad (19)$$

where $\Delta n_f\equiv n_{fy}-n_{fx}$, $\Delta n_R\equiv n_{Ry}-n_{Rx}$ and $\Delta n_r\equiv n_{ry}-n_{rx}$. In the particular case of only two different replicas, one going to x and the other to y and for notational simplicity taking the convention that the most (least) attractive replica goes to y (x), we have $\Delta n_R=1$ and $\Delta n_r=-1$. Therefore,

$$P_y=\left(1+\frac{s_r}{s_R}s_f^{-\Delta n_f}\right)^{-1}. \quad (20)$$

Note that the probability in Eq. **20** does not depend on s_r and s_R separately, but only on their ratio. Therefore, in this case the model uses only two parameters (s_f and s_r/s_R). We compared the model with the stickleback data set from [43], Fig. 6. The data in Fig. 6*B* has a different type of replica pair in each row, so in principle we would fit a different ratio s_r/s_R for each row. But note that the first three rows correspond to experiments with the same three replicas (large, medium and small), combined in different pairs. The same can be

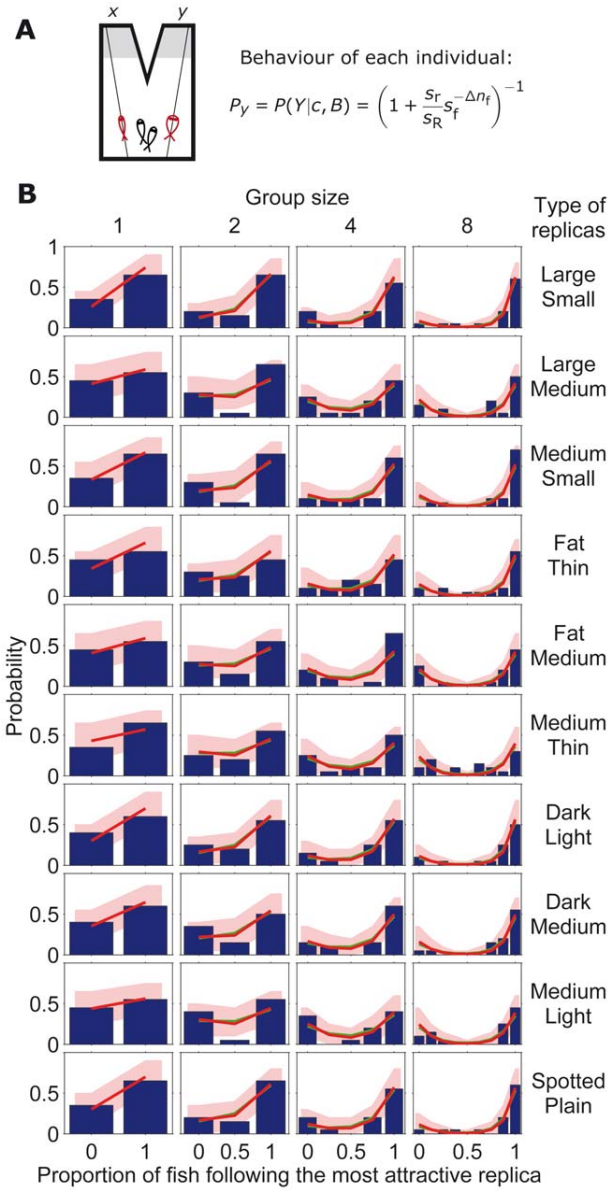


Figure 6. Comparison between model and stickleback choices with two differently modified replicas. (*A*) Schematic diagram of symmetric set-up with a group of sticklebacks (in black) choosing between two identical refugia and with one replica fish going to x and a different one (in size, shape or pattern) going to y (in red). (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [43] (blue histogram) and results from model in Eq. **20** in the main text (red line using reliability parameter $s_f = 2.9$ and $s_r/s_R = 0.35, 0.7, 0.5, 0.52, 0.69, 0.75, 0.43, 0.55, 0.78, 0.43$, for each row from top to bottom; red region: 95% confidence interval; green line with $s_f = 2.5$ and same ratios s_r/s_R as for red line). Different graphs correspond to different stickleback group sizes and different types of replicas going to x and y.
doi:10.1371/journal.pcbi.1002282.g006

said for the second and third threesomes of rows. Therefore, there are only two free parameters for each three rows. On the other hand, s_f should have the same value for all cases. The model again reproduces the experimental results reported in reference [43], obtaining the best fit for $s_f = 2.9$ (Fig. 6*B*). The result is robust, with good fits for $s_f = 2-4$ (Fig. 3, green line) in accord with the value obtained for the case shown in Fig. 2*B*.

Asymmetric set-up

We finally considered the case in which sites x and y are different and the three behaviors are 'go to x' (β_x), 'go to y' (β_y) and 'remain undecided' (β_u). Eq. **12** reduces to

$$P_y = \left(1 + a\, s_x^{n_x} s_y^{n_y} s_u^{N-n_x-n_y}\right)^{-1}. \tag{21}$$

The term $a = P(X|C)/P(Y|C)$ represents the non-social information and in general $a \neq 1$ because the set-up is asymmetric by design. This asymmetry might also affect how a deciding animal takes into account the behaviours of other animals depending on which side they chose, making in general $s_x \neq 1/s_y$. Also, indecision might be informative. For example, if non-social information indicates the possible presence of a predator at y, the indecision of other animals might confirm this to the deciding individual, further biasing the decision towards x. Therefore, we may have $s_u \neq 1$. But it may also be the case that the set-up's asymmetry does not affect the social terms, so we also tested a simpler model in which $s \equiv s_x = 1/s_y$ and $s_u = 1$, giving

$$P_y = \left(1 + a\, s^{-\Delta n}\right)^{-1}. \tag{22}$$

The stickleback dataset reported in reference [42] is ideally suited to test the asymmetric model for the experiments that were performed with a replica predator at the right arm (Fig. 7*A*). The model in Eq. **22** fits best the data with $s = 2.6$ (Fig. 7*B*) and it is robust with a good fit in $s = 2-4$ (Fig. 3, blue line). The more complex model in Eq. **21** gives fits very similar to those of simpler model. Specifically, parameter s_u was rejected by the Bayes Information Criterion [59,60], suggesting that fish do not rely on undecided individuals. The fact that fish rely differently on other fish depending on the option they have taken could not be ruled out by the Bayes Information Criterion, but in any case the impact of this difference on the data is small.

In the experiments in Fig. 2 and Fig. 7, we have assumed that the replicas are perceived by fish as real animals. However, it is reasonable to think that fish might perceive the difference, and rely differently on replicas and real fish. To test this, we considered different behaviors for fish and replicas, such as 'fish goes to x' and 'replica goes to x'. Making that distinction, we get that Eq. **12** reduces to

$$P_y = \left(1 + a\, s_{fx}^{n_{fx}} s_{fy}^{n_{fy}} s_{rx}^{n_{rx}} s_{ry}^{n_{ry}} s_{fu}^{N_f - n_{fx} - n_{fy}}\right)^{-1}. \tag{23}$$

The Bayes Information Criterion rejects only parameter s_{fu}. However, the addition of the new parameters that distinguish replica from real fish give very small improvements in the fits compared to results of the simpler models in Eq. **17** and Eq. **22** (see Fig. S1 and S3), suggesting that fish follow replicas as much as they follow real fish.

Model including dependencies

In this section we will remove the hypothesis of independence among the behaviors of the other individuals (Eq. **6**). We now consider that the focal individual not only takes into account the behaviors of the other animals at the time of decision but the specific sequence of decisions that has taken place before, $\{b_i\}_{i=1}^{K-1}$, being $K-1$ the number of individuals that have decided before the focal one. For example, the sequence $\{x,y\}$ may give different information to the focal individual than the sequence $\{y,x\}$. This

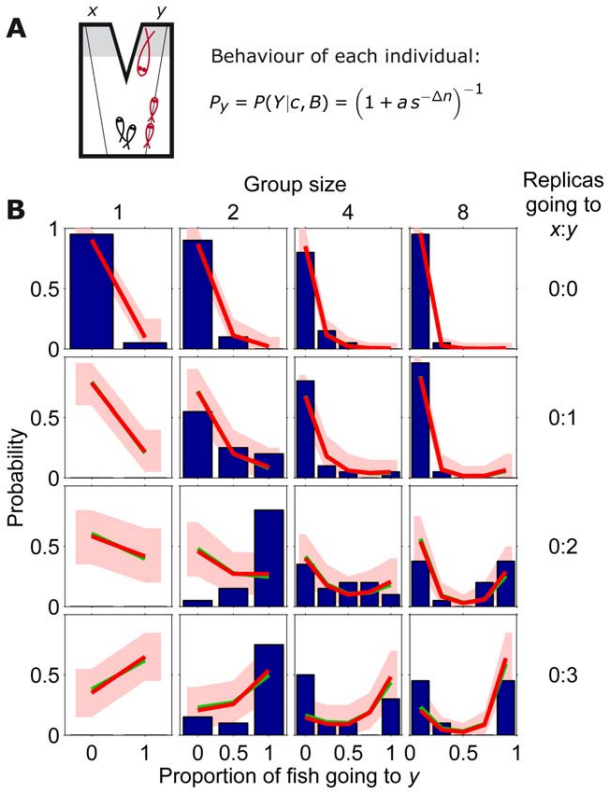


Figure 7. Comparison between model and stickleback choices in asymmetric set-up. (*A*) Schematic diagram of asymmetric set-up (predator at y, large fish depicted in red) with a group of sticklebacks (in black) choosing between two refugia, and replica fish (small fish depicted in red) going to y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from model in Eq. **22** in the main text (red line using $s=2.6$, $a=9.5$; red region: 95% confidence interval. Green line using $s=2.5$ and same a as for red line). Different graphs correspond to different stickleback group sizes and different number of replicas going to y.
doi:10.1371/journal.pcbi.1002282.g007

is illustrated in Fig. 8*A*, where there are two possible paths leading to states labeled as 1:1, but these two states are in different branches of the tree (in contrast with Fig. 4, in which these two states were collapsed in a single one).

To calculate the probability of the observed sequence of behaviors provided that Y is the correct choice,

$$P(\{b_i\}_{i=1}^{K-1}|Y,C_K), \tag{24}$$

one can apply $P(A,B)=P(A|B)P(B)$ repeatedly to obtain

$$P(\{b_i\}_{i=1}^{K-1}|Y,C_K)=\prod_{k=1}^{K-1}P(b_k|Y,C_K,\{b_i\}_{i=1}^{k-1}). \tag{25}$$

This expression substitutes the assumption of independence in Eq. **6**. Each of the terms in the product is simply the probability that the $k-$th individual makes its decision, given the previous decisions, and also given that y is the correct choice. This result was expected since if we look at the tree in Fig. 8*A* we see that the probability of reaching a given state is simply the product of the probabilities of choosing the adequate branches in each step.

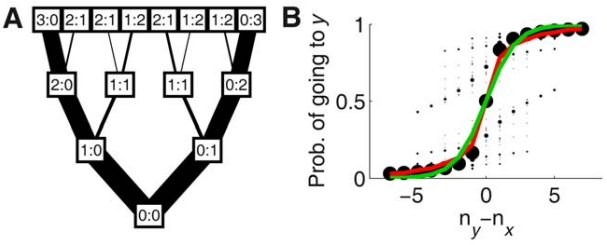


Figure 8. Model taking into account dependencies. (*A*) Decision-making process according to the model with dependencies, Eq. **25**–**33**. Time runs from bottom to top. Each box represents one state, and each edge represents one option of the deciding individual, that either goes to x or to y. Edge width is proportional to the probability of the decision. (*B*) Probability of choosing y as a function of the difference of the number of individuals that have already chosen each option ($\Delta n=n_y-n_x$), for $a'_X=5$. In the new model the probability does not depend any more on Δn alone, so states with the same Δn have different values for the probability (black dots). The area of the dots is proportional to the probability of observing each state. Red line shows the expected value of the probability for each value of Δn. The green line shows the probability for the model that neglects dependencies (Eq. **17**), $\left(1+s^{-\Delta n}\right)^{-1}$ for $s=2.5$.
doi:10.1371/journal.pcbi.1002282.g008

So the problem reduces to computing the individual decision probabilities $P(b_k|Y,C_K,\{b_i\}_{i=1}^{k-1})$. We assume in the following that these probabilities are calculated by the focal individual by assuming that all animals use the same rules to make a decision. The rule for the focal individual is, as in previous sections,

$$P_{y_K}=P(Y|\{b_i\}_{i=1}^{K-1},C_K)=\frac{1}{1+a_K S_K}, \tag{26}$$

where the non-social and social terms are

$$a_K=\frac{P(X|C_K)}{P(Y|C_K)}, \tag{27}$$

and

$$S_K=\frac{P(\{b_i\}_{i=1}^{K-1}|X,C_K)}{P(\{b_i\}_{i=1}^{K-1}|Y,C_K)}, \tag{28}$$

respectively, and where we have added subscript K to S, a and C to reflect that they apply to the focal individual, that makes its decision in the $K-$th place.

The assumption that all animals apply the same rules translates into the following. To apply an equation like Eq. **26** but on a different individual (say, individual k) it is necessary to know the non-social information C_k. Remember that all these computations are made from the point of view of the focal individual, and obviously the focal individual does not have access to the non-social information of the other individuals. It may seem reasonable for the focal animal to assume that all the other individuals have the same non-social information (C_K), but this would result in no social behavior at all (if the other individuals have the same non-social information, their behaviors will not give any extra information). Instead, one can assume that the other individuals may have a different non-social information, C'. Furthermore, this non-social information depends on which is the best choice, because if for example x is the best choice the other individuals have some probability of detecting it, and therefore their non-social information will be on average biased towards x. We

approximate this average bias by assuming that, if x (y) is the best choice, all the other individuals will have non-social information C'_X (C'_Y) that will bias the decision towards x (y). It is therefore the same to assume that x (y) is the best option as to assume that all the other individuals have non-social information C'_X (C'_Y). Therefore, for the probabilities of individual behaviors in Eq. **25**, we have that

$$P(b_k|Y,C_K,\{b_i\}_{i=1}^{k-1}) = P(b_k|C'_Y,C_K,\{b_i\}_{i=1}^{k-1}), \quad (29)$$

where now c'_Y applies to the $k-$th individual, so we can compute this probability simply by applying Eq. **26** to the $k-$th individual,

$$P_{y_k,Y} = \frac{1}{1+a'_Y S_k}, \quad (30)$$

where

$$a'_Y = \frac{P(X|C'_Y)}{P(Y|C'_Y)}. \quad (31)$$

Then, if we denote $P_{b_k,Y} \equiv P(b_k|C'_Y,C_K,\{b_i\}_{i=1}^{k-1})$, we have that

$$\begin{cases} P_{b_k,Y} = P_{y_k,Y} & \text{if } b_k = y \\ P_{b_k,Y} = 1 - P_{y_k,Y} & \text{if } b_k = x. \end{cases} \quad (32)$$

These are the individual probabilities needed in Eq. **25**, that takes into account the correlations among the other individuals. So we can already calculate S_k using Eq. **28**,

$$S_k = \frac{\Pi_{i=1}^{k-1} P_{b_i,X}}{\Pi_{i=1}^{k-1} P_{b_i,Y}}, \quad (33)$$

Eqs. **30** and **33** have a recursive relation, because we need the probabilities up to step $k-1$ to compute S_k, and then we need S_k to compute the probabilities in step k. At the beginning no individual has made any choices, so we start with $S_1 = 1$ and work recursively from there until we obtain the probabilities for individual $K-1$, that allow to compute S_K. Then, we can already use Eq. **26** to compute the decision probability of the focal individual, this time using its actual non-social term a_K (which is 1 for the symmetric cases, and fitted to the data in the non-symmetric case).

The equations above constitute the model taking into account dependencies. The new parameters of this model are a'_X and a'_Y, which substitute s_x and s_y in the previous models, so the number of parameters is exactly the same. In the symmetrical case we must have that $a'_X = 1/a'_Y$, so the model has a single parameter. For the non-symmetrical case these parameters may be independent of each other, but we find good results even assuming that they are not, as was the case for the simplified model. So for simplicity we always assume that

$$a'_X = 1/a'_Y. \quad (34)$$

For the case with different replicas at each side, each of them has a different value of a'_X, thus making one replica more attractive than the other.

The new model also matches very well with the experimental data discussed in this paper. Results for the case of two different replicas are shown in Fig. 9, for the symmetric case in Fig. S4 and for the case with predator in Fig. S5. Fits are robust, and all cases are well explained by the model with the same value of $a'_X = 5$, Fig. S6. See Figs. S1, S2, S3 for a comparison of all models.

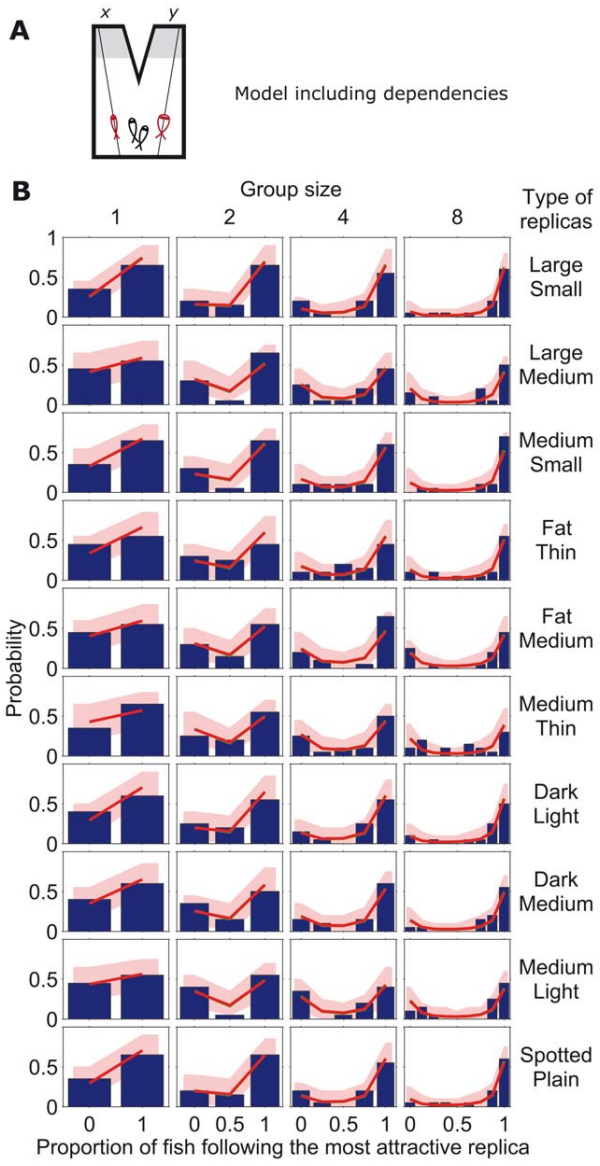


Figure 9. Comparison between model including dependencies and stickleback choices with two differently modified replicas. (*A*) Schematic diagram of symmetric set-up with a group of sticklebacks (in black) choosing between two identical refugia and with one replica fish going to x and a different one (in size, shape or pattern) going to y (in red). (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [43] (blue histogram) and results from model that takes dependencies into account (red line, with $a'_{X,\text{fish}} = 4.8$ and $a'_{X,\text{replicas}} = 21.4, 11.8, 0.6, 9.9, 4.8, 0.9, 13, 8, 0.7, 14.5, 0.9$, for each type of replica (large, medium, small, fat, etc.); red region: 95% confidence interval; green line with $a'_{X,\text{fish}} = 5$ and same $a'_{X,\text{replicas}}$ as for red line). Different graphs correspond to different stickleback group sizes and different types of replicas going to x and y.
doi:10.1371/journal.pcbi.1002282.g009

We now ask how different is the model including dependencies from the model that neglects them. To compare the two models, we plot the probability of going to y as a function of $\Delta n = n_y - n_x$ for the new model, as we did in Fig. 1*B* for the old one. The

inclusion of dependencies has the consequence that the probability of going to y does not depend only on Δn, since now different states with the same Δn may have different probabilities. Therefore, when we plot the probability of going to y as a function of Δn we obtain different values of the probability for each value of Δn. This is shown by the black dots in Fig. 8*B*, where the size of the dots is proportional to the probability of observing each state when starting from 0:0. The red line shows the average probability for each Δn, taking into account the probability of each state. Both the dots and this line correspond to $a'_X = 5$, which is the one that fits best the data. The green line corresponds to the probability for the simplest model neglecting dependencies, with the value that best fits to the data ($s = 2.5$). This line is close to the mean probability for the new model and to the values with highest probability of occurrence, so the simple model is as a good approximation to the model with dependencies.

We find an interesting prediction of the new model: There are some states in which the most likely option is to choose the option chosen by *fewer* individuals (for example, note in Fig. 8*D* that some points with $\Delta n < 0$ are above 0.5). This surprising result comes from the fact that, as more fish accumulate at one side, their choices become less and less informative (because it is very likely that they are simply following the others). If then one fish goes to the opposite side, its behavior is very informative, because it is contradicting its social information. This effect can be so strong that it may beat the effect of all the other individuals, resulting in a higher probability of following this last individual than all the individuals that decided before.

Discussion

We have shown that probabilistic estimation in the presence of uncertainty can explain collective animal decisions. This approach generated a new expression for each experimental manipulation, Eq. **17**–**22**, and was naturally extended to test for more refined cognitive capacities, Eq. **23**. The model was found to have a good correspondence with the data in three experimental settings (Figs. 2, 6 and 7), always giving a good fit with the social reliability parameter s in the interval 2–4. Indeed, all the data have a very good fit with $s = 2.5$ (Figs. 2, 6 and 7, green lines). According to Eq. **9**, this value for s has the interpretation that, for the behaviors relevant for these experiments, the fish assume that their conspecifics make the right choice 2.5 times more often than the wrong choice.

For the data used in this paper, previous empirical fits used more parameters [42] (Figs. S1, S2, S3, blue line), and added more complex behavioral rules when the basic model failed [43] (Fig. S2, blue line). Our approach thus gains in simplicity. It also finds an expression for each set-up with expressions for complex set-ups obtained with add-ons to those of simpler set-ups, making the model scalable and easier to understand in terms of simpler experiments. Also, taking the models as fits to experimental data, the bayesian information criterion finds our models to be better than those in [42] and [43] (see captions in Figs. S1, S2, S3 for details).

Collective animal behavior has been subject to a particularly careful quantitative analysis. Previous studies have given descriptions led by the powerful idea that complex collective behaviors can emerge from simple individual rules. In fact, some systems have been found empirically to obey rules that are mathematically similar or the same as some of the ones presented in this paper, further supporting the idea that probabilistic estimation might underlie collective decision rules in many species. For example, a function like the one in Eq. **17** has been used to describe the behavior of Pharaoh's ant [61], a function like Eq. **22** for mosquito fish [62], and a function like the one in the right-hand-side of Eq. **22** for meerkats [63]. But despite the importance of group decisions in animals, little is known about the origin of such simple individual rules. This paper argues that probabilistic estimation can be an underlying substrate for the rules explaining collective decisions, thus helping in their evolutionary explanation. Also, this connection between patterns in animal collectives and a cognitive process helps to explain the similarities that exist between decision-making processes at the level of the brain and at the level of animal collectives [64,65].

Our model is naturally compatible with other theories that use a Bayesian formalism to study different aspects of behavior and neurobiology, thus contributing to a unified approach of information processing in animals. For example, it may be combined with the formalism of Bayesian foraging theory [18], through an expansion of the non-social reliability a. Related to this case, a very well studied example of use of social information is the one in which one individual can observe directly the food collected by another individual [29–33]. In this case the social information is as unambiguous as the non-social one, so in this case both types of information should have a similar mathematical form [29–33]. This is consistent with our model, that in this case will give a similar expression for a and S. Other kinds of social information (such as another individual's decision to leave a food patch or choices of females in mating [41]) would enter naturally in our reliability terms s_k. In discussing these and similar problems, it has been proposed that animals should use social information when their personal information is poor, and ignore it otherwise [25,26,41]. Our model provides a quantitative framework for this problem, predicting that social information is always used, only with different weights with respect to other sources of information. Bayesian estimation is also a prominent approach to study decisions in neurobiology and psychology [3–17] and it would be of interest to explore the mechanisms and role played by the multiplicative relation between non-social and social terms.

Our approach also makes a number of predictions. For example, it derives the probability of choosing among M options (see Eq. **S16** of the *Text S1*), that for the symmetric case reduces to

$$P_\mu = \left(1 + \sum_{m \neq \mu}^{M} s^{-(n_\mu - n_m)}\right)^{-1}, \tag{35}$$

predicted also to fit the data for cases with $M > 2$ options.

We also predict a quantitative link between estimation and collective behavior. The parameters a and s_k in our model are in fact not merely fitting parameters, but true experimental variables. Manipulations of a and s_k should allow to test that changes in collective behavior follow the predictions of the model. A counterintuitive prediction about the manipulation of s_k is that external factors unrelated to the social component can nevertheless modify it. For example, a fish that usually finds food in a given environment should interpret a sudden turn of one of his mates as an indication that it has found food, and therefore will follow it. In contrast, another fish that is not expected to find food in that environment will not interpret the sudden turn as indicative of food, and will not follow. Thus, the model predicts that the *a priori* probability of finding food (to which each fish can be trained in isolation) will modify its propensity to follow conspecifics. An alternative approach that would not need manipulation of the reliabilities s_k would consist in showing that the probability of copying a behavior increases with how reliably the behavior informs about the environment.

We can also extend the estimation model to use, instead of the location of animals, their predicted location. We would then find expressions like the ones in this paper but for the number or density of individuals estimated for a later time. Consider for example the case without non-social information, described in Eq. **17** for two options and in Eq. **35** for more options. We can rewrite these equations as $P_\mu = \Omega s^{n_\mu}$ with μ one of the options and Ω is the normalization, $\Omega = \sum_{m=1}^{M} s^{n_m}$, where M is the number of options. Then, we would have $P(\vec{x}) = \Omega s^{\rho(\vec{x};t+\Delta t)}$ for the continuous case using prediction. Future positions at times $t+\Delta t$ (where Δt does not need to be constant) in terms of variables at present time t would be given by $\vec{x}+\vec{v}\Delta t$ for animals moving at constant velocity $\vec{v}$. Consider then a simple case of an animal located at $\vec{x}$ and estimating the future position of a compact group at $\vec{x}_g$ and moving with velocity $\vec{v}_g$. The deciding animal would be predicted to move with a high probability in the direction $\left(\vec{x}_g(t)-\vec{x}(t)\right)+\Delta t\vec{v}_g(t)$. Estimation of future locations thus naturally predicts in this simple case a particular form of 'attraction' and 'alignment' forces of dynamical empirical models [46,66] as attraction to future positions, but in the general also deviations from these simple rules.

Methods

Obtaining group behavior from the model of an individual

The estimation rules presented in this paper refer to a single individual. To simulate the behavior of a group, we use the following algorithm: The current individual decides between x and y. After the decision, we recompute the relevant parameters of the model and use the new values for the next deciding individual. The undecided individuals are only those that are waiting for their turn to decide. We tested an alternative implementation in which individuals may remain undecided or in which two individuals can decide simultaneously, obtaining no relevant differences.

For the case of the model including dependencies, the model always starts at state 0:0, with $S=1$. Most experiments have initial conditions in which several replicas are already going to either side, and the fish have no information about the path followed to reach this state. In these cases, we average the probabilities of all the paths that might have possibly led to the initial state to compute the initial value of S.

Protocol S1 and *Protocol S2*, contain Matlab functions that run the models (extensions of the files must be changed from .txt to .m to make them operative). *Protocol S1* corresponds to the model without dependencies, and *Protocol S2* corresponds to the model with dependencies. These functions have been used to generate all the theoretical results presented in this paper.

Fits

We computed log likelihood as the logarithm of the probability that the histograms come from the model. We searched for the model parameters giving a higher value of log likelihood, corresponding to a better fit. This search was performed by optimizing each parameter separately (keeping the rest constant) and iterating through all parameters until convergence. In all cases convergence was rapidly achieved. We performed multiple searches for best fitting parameters starting from random initial conditions and always found convergence to the same values, suggesting there are no local maxima. Indeed, we observed that log-likelihood is smooth and with a single maximum in all the cases with 1 or 2 parameters (see Fig. 3 for an example).

Bayesian Information Criterion

For model comparison we used the Bayesian Information Criterion (BIC) [59,60], which takes into account both goodness of fit and the number of parameters. According to this criterion, among several models that have been fitted to maximize log likelihood, one should select the one for which

$$BIC_i = L_i - \frac{1}{2}k_i \log(h) \tag{36}$$

is largest, where L_i is the logarithm of the probability that the data comes from the $i-$th model once its parameters have been optimized to maximize this probability, k_i is its number of parameters of the $i-$th model and h is the number of measurements (which in our case is the same for all models).

More intuitive than the direct BIC_i values in Eq. **36** are the BIC weights, defined as [60]

$$w_i = \frac{\exp(BIC_i)}{\sum_j \exp(BIC_j)}, \tag{37}$$

when we assume that all models are *a priori* equally likely. Roughly speaking, w_i can be interpreted as the probability that model i is the most correct one [60].

We used BIC to compare different versions of our model, and also to compare our model with those of references [42,43] (see Figs. S1, S2, S3). The models of refs. [42,43] were originally fitted by minimizing the mean squared error instead of by maximizing logprob. For this reason, they score very poorly in BIC with their reported parameters. For this reason, we re-optimized for maximum logprob all their model parameters (these parameters are, using the notation of refs. [42,43], a, k, T, m and r, with r only applicable in the case of predator present). For the case of different replicas going to each side, parameter p_{bias} takes a different value for each row in the figure, adding up to 10 parameters. The model in ref. [43] is computationally expensive, so it is not feasible to re-optimize these many parameters. Therefore, we treated them as if they were independently measured: we fixed p_{bias} in each case so that the results of the trials with a single individual matched exactly the model's prediction (as reported in [43]). We also followed this procedure with the ratios s_r/s_R of our model without dependencies, and the pairs $a'_{X,\text{replicas}}$ in our model with dependencies. Then, we performed BIC taking into account neither these parameters (p_{bias} the ratios s_r/s_R and the pairs $a'_{X,\text{replicas}}$) nor the data from trials using single individuals.

Supporting Information

Figure S1 Comparison between different models for the symmetric set-up. Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histograms). Red line: results from our single-parameter model assuming independence in Eq. **17** in the main text ($s=2.2$). Green line: Enhanced model assuming independence with different reliability for the replicas ($s_f=3$, $s_r=1.76$). Yellow line: Model including dependencies ($a'_X=4.9$). Blue line: Empirical model presented in Ref. [42], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different number of replicas going to x and y. According to Bayesian Information Criterion (BIC, see *Methods*), the best model is our model with dependencies (yellow line, logprob $L=-394$, and BIC weight $w=0.996$. Second-best is the complicated version

of the model without dependencies (green line, logprob $L=-396$, and BIC weight $w=0.004$). Third-best is our one-parameter model assuming independence (red line, $L=-419$, $w=3\cdot10^{-11}$). And last (but not far from the third one) the model from Ref. [42] (blue line, $L=-411$ $w=5\cdot10^{-13}$). For the model from Ref. [42], L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [42] would perform worse).
(TIF)

Figure S2 Comparison between different models for the condition with two different replicas. Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [43] (blue histograms). Red line: results from model in Eq. **20** in the main text ($s_f=2.9$, $s_r/s_R=0.35$, 0.7, 0.5, 0.52, 0.69, 0.75, 0.43, 0.55, 0.78, 0.43 for each row from top to bottom). Yellow line: Model including dependencies ($a'_X=4.8$, $a'_{X,\text{replicas}}=21.4$, 11.8, 0.6, 9.9, 4.8, 0.9, 13, 8, 0.7, 14.5, 0.9 for each type of replica (large, medium, small, etc.). Blue line: Empirical model presented in Ref. [43], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different types of replicas going to x and y. According to Bayesian Information Criterion (BIC, see *Methods*), our model neglecting dependencies gives the best representation of the data (red line, logprob $L=-783$, and BIC weight $w=0.9985$). Second-best is out model including dependencies, ($L=-788$, $w=0.001$). Last, but near the second one, is the model from ref. [43] (blue line, $L=-781$ $w=0.0005$. For the model from Ref. [43], these values of L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [43] would perform worse). The values of logprob (L) reported here do not include the data of the single-individual experiments (see *Methods*).
(TIF)

Figure S3 Comparison between different models in the asymmetrical set-up. Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histograms). Red line: results from model neglecting dependencies in Eq. **22** in the main text ($s=2.6$, $a=9.5$). Green line: Enhanced model neglecting dependencies with different reliability for the fish going to different locations and for the replicas ($a=5.5$, $s_{fx}=50$, $s_{fy}=2/3$, $s_{ry}=0.36$. s_{rx} has no effect because there are no replicas going to x). Yellow line: Two-parameter model including dependencies ($a=9.94$, $a'_X=8.66$). Blue line: Empirical model presented in Ref. [42], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different number of replicas going to y. According to Bayesian Information Criterion (BIC, see *Methods*), the best two models are our complicated version neglecting dependencies (green line, logprob $L=-225$, and BIC weight $w=0.52$) and our two-parameter model including dependencies (yellow line, $L=-231$, $w=0.38$). Next (but very near) is our simplified model (red line, $L=-232$, $w=0.098$). And last (and significantly worse) the model from Ref. [42] (blue line, $L=-234$ $w=2.5\cdot10^{-6}$. For the model from Ref. [42], the values of L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [42] would perform worse. In two of the graphs for group size 1 that there are no data the prediction of the model from Ref. [42] and our model (especially the simplest version) are opposite. It might be that the results changed completely, depending on the results of these graphs, were the experiments performed. But we found that this is not the case: We performed simulations, adding experimental data in these two graphs. Even in the extreme case that the fabricated results matched exactly the predictions of the model in Ref. [42], BIC would still favour two of our models (we would get $L=-254$, $w=0.99$ for our model with dependence, $L=-252$, $w=0.01$ for our complicated model neglecting dependence, $L=-268$, $w=8\cdot10^{-7}$ for our simplified model neglecting dependence and $L=-258$, $w=3\cdot10^{-6}$ for the model in [42]).
(TIF)

Figure S4 Comparison between model including dependencies and stickleback choices in symmetric set-up. (*A*) Schematic diagram of symmetric set-up with a group of sticklebacks (in black) choosing between two identical refugia and with different numbers of replica fish (in red) going to x and y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from the model that takes into account dependencies (red line using $a'_X=4.9$; red region: 95% confidence interval; green line with $a'_X=5$). Different graphs correspond to different stickleback group sizes and different number of replicas going to x and y.
(TIF)

Figure S5 Comparison between model including dependencies and stickleback choices in asymmetric set-up. *A*) Schematic diagram of asymmetric set-up (predator at y, large fish depicted in red) with a group of sticklebacks (in black) choosing between two refugia, and replica fish (small fish depicted in red) going to y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from the model that takes into account the dependencies (red line using $a'_X=8.7$, $a=9.9$; red region: 95% confidence interval. Green line using $a'_X=5$ and $a=6.28$). Different graphs correspond to different stickleback group sizes and different number of replicas going to y.
(TIF)

Figure S6 Goodness of fit of the model including dependencies for different values of a'_X. Red: Symmetric case (data in Fig. S4). **Green:** Case with different replicas at each side (data in Fig. 9. The parameters $a'_{X,\text{replicas}}$ are re-optimized for each value of a'_X). **Blue:** Asymmetric set-up with predator on one side (data in Fig. S5; Parameter a is re-optimized for each value of a'_X). (*A*) Root mean squared error between the data and the probabilities predicted by the model. Grey dashed line shows the mean RMSE for the three cases. The absolute values for each case depend on the shape of the data and are not comparable, only the trends and the position of the minima should be compared. (*B*) Logarithm of the probability that the data come from the model. The height of each curve depends on the number of data for each experiment, only the trend and the position of the maxima should be compared. Grey dashed line shows the sum of the three coloured lines, but shifted by 1000 so that it fits on the scale. The peak of this global probability indicates the value of a'_X that best fits the three datasets ($a'_X=5$).
(TIF)

Protocol S1 Algorithm for the model that neglects dependencies. This file contains Matlab code that runs the model without dependencies. Please, change extension from .txt to .m to make it operative. It can be run without any input argument. Once the extension is changed to .m, simply type ProtocolS1 in Matlab's command window to get results for default parameters. Documentation is given inside the file. Type help ProtocolS1 in Matlab's command window to see the documentation.
(TXT)

Protocol S2 Algorithm for the model that takes dependencies into account. This file contains Matlab code that runs the model with dependencies. Please, change extension from .txt to .m to make it operative. It can be run without any input argument. Once the extension is changed to .m, simply type ProtocolS2 in Matlab's command window to get results for default parameters. Documentation is given inside the file. Type help ProtocolS2 in Matlab's command window to see the documentation.
(TXT)

Text S1 Derivation of the model with more options. This file contains the derivation of the model for the more general case of M different options (instead of only 2, as presented in the main text).
(PDF)

Acknowledgments

We acknowledge useful comments by Sara Arganda, Larissa Conradt, Iain Couzin, Jacques Gautrais, David Sumpter, Guy Theraulaz, Julián Vicente Page and COLMOT 2010 participants.

Author Contributions

Conceived and designed the experiments: APE GGdP. Performed the experiments: APE GGdP. Analyzed the data: APE GGdP. Wrote the paper: APE GGdP.

References

1. Box G, Tiao G (1973) Bayesian inference in statistical analysis. New York: Addison-Wesley, Available: http://onlinelibrary.wiley.com/doi/10.1002/9781118033197.fmatter/summary.
2. Jaynes ET, Bretthorst LG (2003) Probability Theory: The Logic of Science (Vol 1) Cambridge University Press.
3. Helmholtz H (1925) Physiological Optics, Vol. III: The perceptions of Vision. RochesterNY, USA: Optical Society of America.
4. Mach E (1980) Contributions to the Analysis of the Sensations. ChicagoIL, USA: Open Court Publishing Co.
5. Knill DC, Pouget A (2004) The Bayesian brain: the role of uncertainty in neural coding and computation. Trends Neurosci 27: 712–9.
6. Jacobs R (1999) Optimal integration of texture and motion cues to depth. Vision Res 39: 3621–3629.
7. Knill DC, Saunders JA (2003) Do humans optimally integrate stereo and texture information for judgments of surface slant? Vision Res 43: 2539–2558.
8. Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415: 429–33.
9. Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. J Opt Soc Am A 20: 1391.
10. Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. Curr Biol 14: 257–262.
11. Gold JI, Shadlen MN (2001) Neural computations that underlie decisions about sensory stimuli. Trends Cogn Sci 5: 10–16.
12. Kording KP, Wolpert DM (2004) Bayesian integration in sensorimotor learning. Nature 427: 244–247.
13. Körding KP, Wolpert DM (2006) Bayesian decision theory in sensorimotor control. Trends Cogn Sci 10: 319–26.
14. Gold JI, Shadlen MN (2007) The neural basis of decision making. Annu Rev Neurosci 30: 535–74.
15. Courville AC, Daw ND, Touretzky DS (2006) Bayesian theories of conditioning in a changing world. Trends Cogn Sci 10: 294–300.
16. Kruschke JK (2006) Locally Bayesian learning with applications to retrospective revaluation and highlighting. Psychol Rev 113: 677–99.
17. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND (2011) How to Grow a Mind: Statistics, Structure, and Abstraction. Science 331: 1279–1285.
18. Oaten A (1977) Optimal foraging in patches: A case for stochasticity. Theor Popul Biol 12: 263–285.
19. Biernaskie JM, Walker SC, Gegear RJ (2009) Bumblebees learn to forage like Bayesians. Am Nat 174: 413–423.
20. Alonso J (1995) Patch use in cranes: a field test of optimal foraging predictions. Anim Behav 49: 1367–1379.
21. McNamara JM, Green RF, Olsson O (2006) Bayes theorem and its applications in animal behaviour. Oikos 112: 243–251.
22. Valone TJ (2006) Are animals capable of Bayesian updating? An empirical review. Oikos 112: 252–259.
23. Valone TJ, Templeton JJ (2002) Public information for the assessment of quality: a widespread social phenomenon. Philos Trans R Soc Lond B Biol Sci 357: 1549–57.
24. Blanchet S, Clobert J, Danchin E (2010) The role of public information in ecology and conservation: an emphasis on inadvertent social information. Ann NY Acad Sci 1195: 149–68.
25. Dall SRX, Giraldeau LA, Olsson O, McNamara JM, Stephens DW (2005) Information and its use by animals in evolutionary ecology. Trends Ecol Evol 20: 187–93.
26. Giraldeau LA, Valone TJ, Templeton JJ (2002) Potential disadvantages of using socially acquired information. Philos Trans R Soc Lond B Biol Sci 357: 1559–66.
27. Wagner RH, Danchin E (2010) A taxonomy of biological information. Oikos 119: 203–209.
28. King AJ, Cowlishaw G (2007) When to use social information: the advantage of large group size in individual decision making. Biol Lett 3: 137–9.
29. Valone TJ (1989) Group Foraging, Public Information, and Patch Estimation. Oikos 56: 357–363.
30. Templeton JJ, Giraldeau LA (1995) Patch assessment in foraging flocks of European starlings: Evidence for the use of public information. Behav Ecol 6: 65–72.
31. Templeton JJ, Giraldeau LA (1996) Vicarious sampling: The use of personal and public information by starlings foraging in a simple patchy environment. Behav Ecol Sociobiol 38: 105–14.
32. Smith JW, Benkman CW, Coffey K (1999) The use and misuse of public information by foraging red crossbills. Behav Ecol 10: 54–62.
33. Clark C, Mangel M (1986) The evolutionary advantages of group foraging. Theor Popul Biol 30: 45–75.
34. Doligez B, Danchin E, Clobert J (2002) Public information and breeding habitat selection in a wild bird population. Science 297: 1168–70.
35. Boulinier T, Danchin E (1997) The use of conspecific reproductive success for breeding patch selection in terrestrial migratory species. Evol Ecol 11: 505–517.
36. Coolen I, van Bergen Y, Day RL, Laland KN (2003) Species difference in adaptive use of public information in sticklebacks. Proc Biol Sci 270: 2413–9.
37. van Bergen Y, Coolen I, Laland KN (2004) Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. Proc Biol Sci 271: 957–62.
38. Rieucau G, Giraldeau La (2009) Persuasive companions can be wrong: the use of misleading social information in nutmeg mannikins. Behav Ecol 20: 1217–1222.
39. Lima SL (1995) Collective detection of predatory attack by social foragers: fraught with ambiguity? Anim Behav 50: 1097–1108.
40. Proctor CJ, Broom M, Ruxton GD (2001) Modelling antipredator vigilance and flight response in group foragers when warning signals are ambiguous. J Theor Biol 211: 409–17.
41. Nordell, Valone TJ (1998) Mate choice copying as public information. Ecol Lett 1: 74–76.
42. Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J (2008) Quorum decision-making facilitates information transfer in fish shoals. Proc Natl Acad Sci USA 105: 6948–53.
43. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. Curr Biol 18: 1773–1777.
44. Couzin ID, Krause J (2003) Self-organization and collective behavior in vertebrates. Adv Stud Behav 32: 1–75.
45. Sumpter DJ (2006) The principles of collective animal behaviour. Philos Trans R Soc Lond B Biol Sci 361: 5–22.
46. Couzin ID, Krause J, Franks NR, Levin SA (2005) Effective leadership and decision-making in animal groups on the move. Nature 433: 513–516.
47. Katz Y, Tunstrom K, Ioannou CC, Huepe C, Couzin ID (2011) Inferring the structure and dynamics of interactions in schooling fish. Proc Natl Acad Sci USA, E-pub ahead of print.
48. Neyman J, Pearson E (1933) On the problem of the most efficient tests of statistical hypotheses. Philos Transact A Math Phys Eng Sci 231: 289.
49. Herrnstein R (1961) Relative and absolute strength of response as a function of frequency of reinforcement. J Exp Anal Behav 4: 267.
50. Behrend ER, Bitterman ME (1961) Probability-Matching in the Fish. Am J Psychol 74: 542–551.
51. Greggers U, Menzel R (1993) Memory dynamics and foraging strategies of honeybees. Behav Ecol Sociobiol 32: 17–29.
52. Kirk KL, Bitterman ME (1965) Probability-Learning by the Turtle. Science 148: 1484–1485.
53. Vulkan N (2000) An Economist's Perspective on Probability Matching. J Econ Surv 14: 101–118.
54. Wozny DR, Beierholm UR, Shams L (2010) Probability matching as a computational strategy used in perception. PLoS Comput Biol 6: 7.
55. Staddon J (1983) Adaptive Behavior and Learning. Cambridge: Cambridge University Press, Available: http://dukespace.lib.duke.edu/dspace/handle/10161/2878.
56. Fretwell S, Lucas H (1969) On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheor 19: 16–36.

57. Houston A, McNamara J (1987) Switching between resources and the ideal free distribution. Anim Behav 35: 301–302.
58. Gaissmaier W, Schooler LJ (2008) The smart potential behind probability matching. Cognition 109: 416–22.
59. Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6: 461–464.
60. Link WA, Barker RJ (2006) Model weights and the foundations of multimodel inference. Ecology 87: 2626–2635.
61. Jeanson R, Ratnieks FLW, Deneubourg JL (2003) Pheromone trail decay rates on different substrates in the Pharaoh's ant, Monomorium pharaonis. Physiol Entomol 28: 192–198.
62. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. Proc Natl Acad Sci USA 108: 6–9.
63. Bousquet CAH, Sumpter DJT, Manser MB (2011) Moving calls: a vocal mechanism underlying quorum decisions in cohesive groups. Proc Biol Sci 278: 1482–1488.
64. Marshall JA, Bogacz R, Dornhaus A, Planqué R, Kovacs T, et al. (2009) On optimal decisionmaking in brains and social insect colonies. J Roy Soc Interface 6: 1065–74.
65. Couzin ID (2009) Collective cognition in animal groups. Trends Cogn Sci 13: 36–43.
66. Couzin ID, Krause J, James R, Ruxton GD, Franks NR (2010) Collective memory and spatial sorting in animal groups. J Theor Biol 218: 1–11.

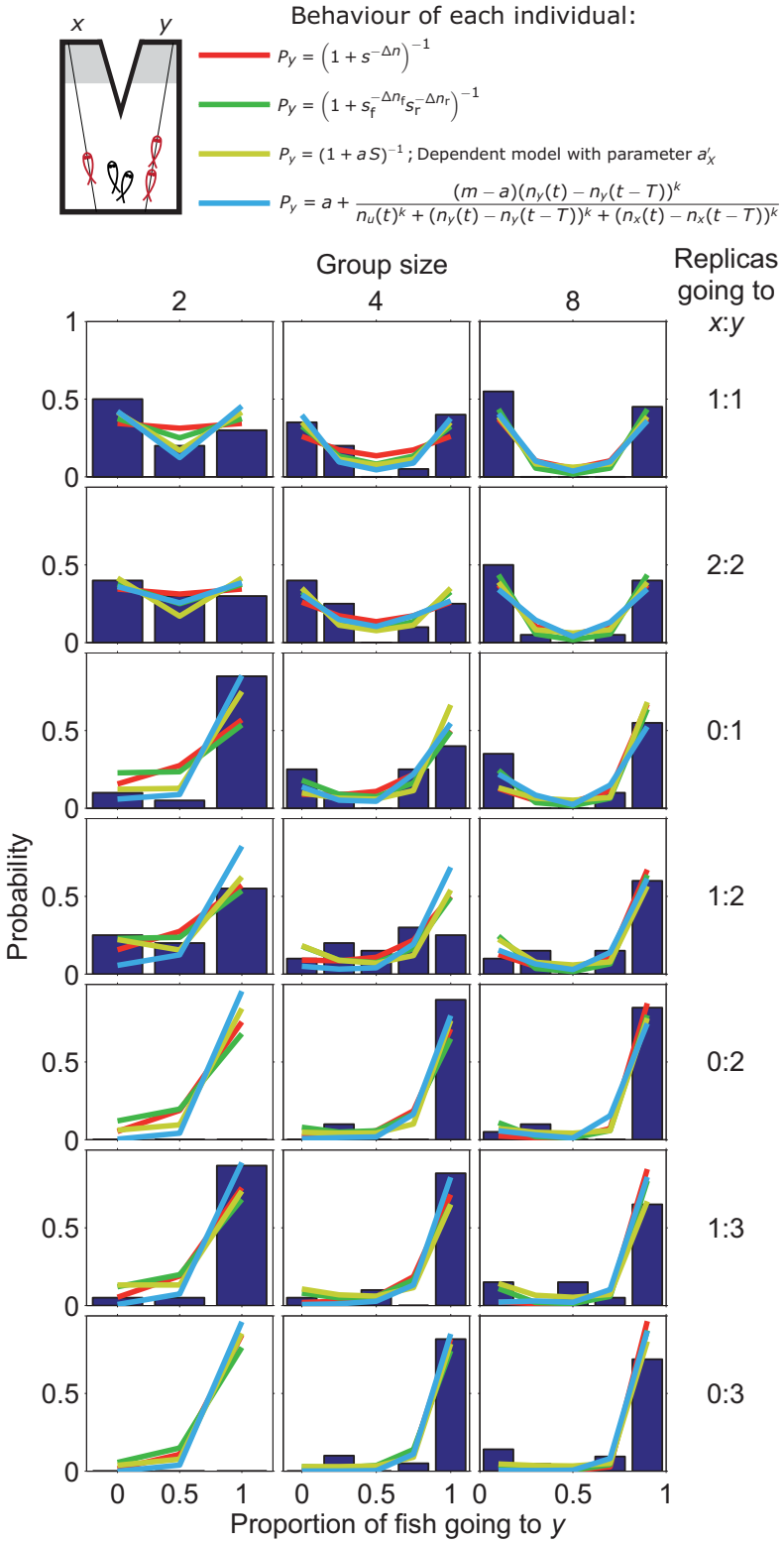


Figure S1: *Comparison between different models for the symmetric set-up.* Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histograms). Red line: results from our single-parameter model assuming independence in Eq. 17 in the main text ($s = 2.2$). Green line: Enhanced model assuming independence with different reliability for the replicas ($s_f = 3$, $s_r = 1.76$). Yellow line: Model including dependencies ($a'_X = 4.9$). Blue line: Empirical model presented in Ref. [42], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different number of replicas going to x and y. According to Bayesian Information Criterion (BIC, see *Methods*), the best model is our model with dependencies (yellow line, logprob $L = -394$, and BIC weight $w = 0.996$. Second-best is the complicated version of the model without dependencies (green line, logprob $L = -396$, and BIC weight $w = 0.004$). Third-best is our one-parameter model assuming independence (red line, $L = -419$, $w = 3 \cdot 10^{-11}$). And last (but not far from the third one) the model from Ref. [42] (blue line, $L = -411$ $w = 5 \cdot 10^{-13}$). For the model from Ref. [42], L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [42] would perform worse).

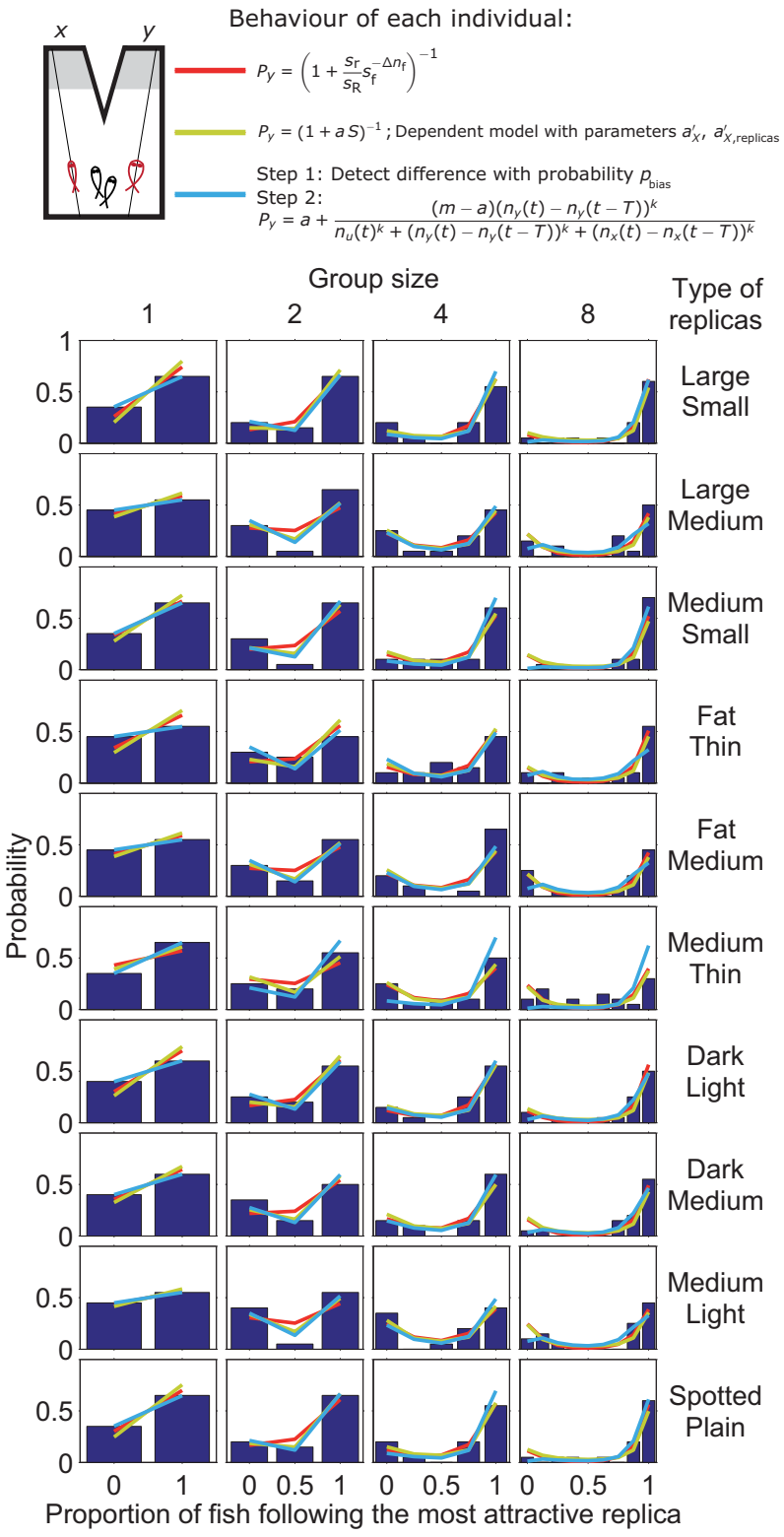


Figure S2: *Comparison between different models for the condition with two different replicas.* Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [43] (blue histograms). Red line: results from model in Eq. 20 in the main text ($s_f = 2.9$, s_r/s_R = 0.35, 0.7, 0.5, 0.52, 0.69, 0.75, 0.43, 0.55, 0.78, 0.43 for each row from top to bottom). Yellow line: Model including dependencies (a'_X = 4.8, $a'_{X,\text{replicas}}$ = 21.4, 11.8, 0.6, 9.9, 4.8, 0.9, 13, 8, 0.7, 14.5, 0.9 for each type of replica (large, medium, small, etc.). Blue line: Empirical model presented in Ref. [43], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different types of replicas going to x and y. According to Bayesian Information Criterion (BIC, see *Methods*), our model neglecting dependencies gives the best representation of the data (red line, logprob $L = -783$, and BIC weight $w = 0.9985$). Second-best is out model including dependencies, ($L = -788$, $w = 0.001$). Last, but near the second one, is the model from ref. [43] (blue line, $L = -781$ $w = 0.0005$. For the model from Ref. [43], these values of L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [43] would perform worse). The values of logprob (L) reported here do not include the data of the single-individual experiments (see *Methods*).

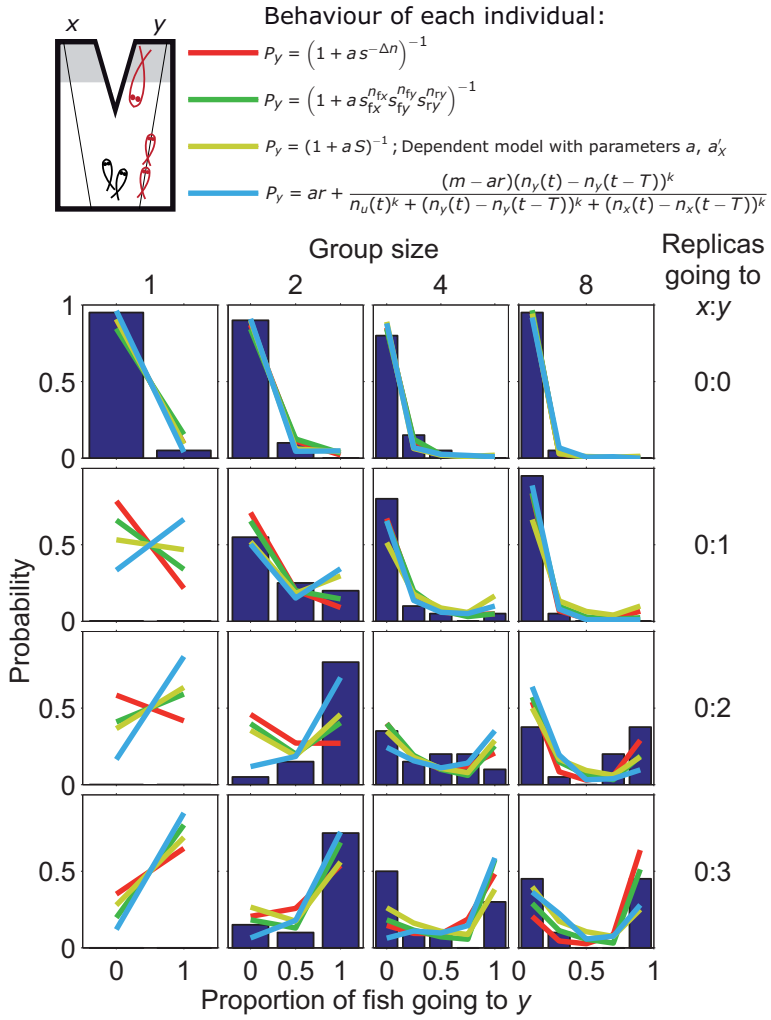


Figure S3: *Comparison between different models in the asymmetrical set-up.* Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histograms). Red line: results from model neglecting dependencies in Eq. 22 in the main text ($s = 2.6$, $a = 9.5$). Green line: Enhanced model neglecting dependencies with different reliability for the fish going to different locations and for the replicas ($a = 5.5$, $s_{fx} = 50$, $s_{fy} = 2/3$, $s_{ry} = 0.36$. s_{rx} has no effect because there are no replicas going to x). Yellow line: Two-parameter model including dependencies ($a = 9.94$, $a'_X = 8.66$). Blue line: Empirical model presented in Ref. [42], using the parameters reported there. Different graphs correspond to different stickleback group sizes and different number of replicas going to y. According to Bayesian Information Criterion (BIC, see *Methods*), the best two models are our complicated version neglecting dependencies (green line, logprob $L = -225$, and BIC weight $w = 0.52$) and our two-parameter model including dependencies (yellow line, $L = -231$, $w = 0.38$). Next (but very near) is our simplified model (red line, $L = -232$, $w = 0.098$). And last (and significantly worse) the model from Ref. [42] (blue line, $L = -234$ $w = 2.5 \cdot 10^{-6}$. For the model from Ref. [42], the values of L and w correspond to a re-optimization of the model as described in *Methods*, because using the parameters reported in [42] would perform worse. In two of the graphs for group size 1 that there are no data the prediction of the model from Ref. [42] and our model (especially the simplest version) are opposite. It might be that the results changed completely, depending on the results of these graphs, were the experiments performed. But we found that this is not the case: We performed simulations, adding experimental data in these two graphs. Even in the extreme case that the fabricated results matched exactly the predictions of the model in Ref. [42], BIC would still favour two of our models (we would get $L = -254$, $w = 0.99$ for our model with dependence, $L = -252$, $w = 0.01$ for our complicated model neglecting dependence, $L = -268$, $w = 8 \cdot 10^{-7}$ for our simplified model neglecting dependence and $L = -258$, $w = 3 \cdot 10^{-6}$ for the model in [42]).

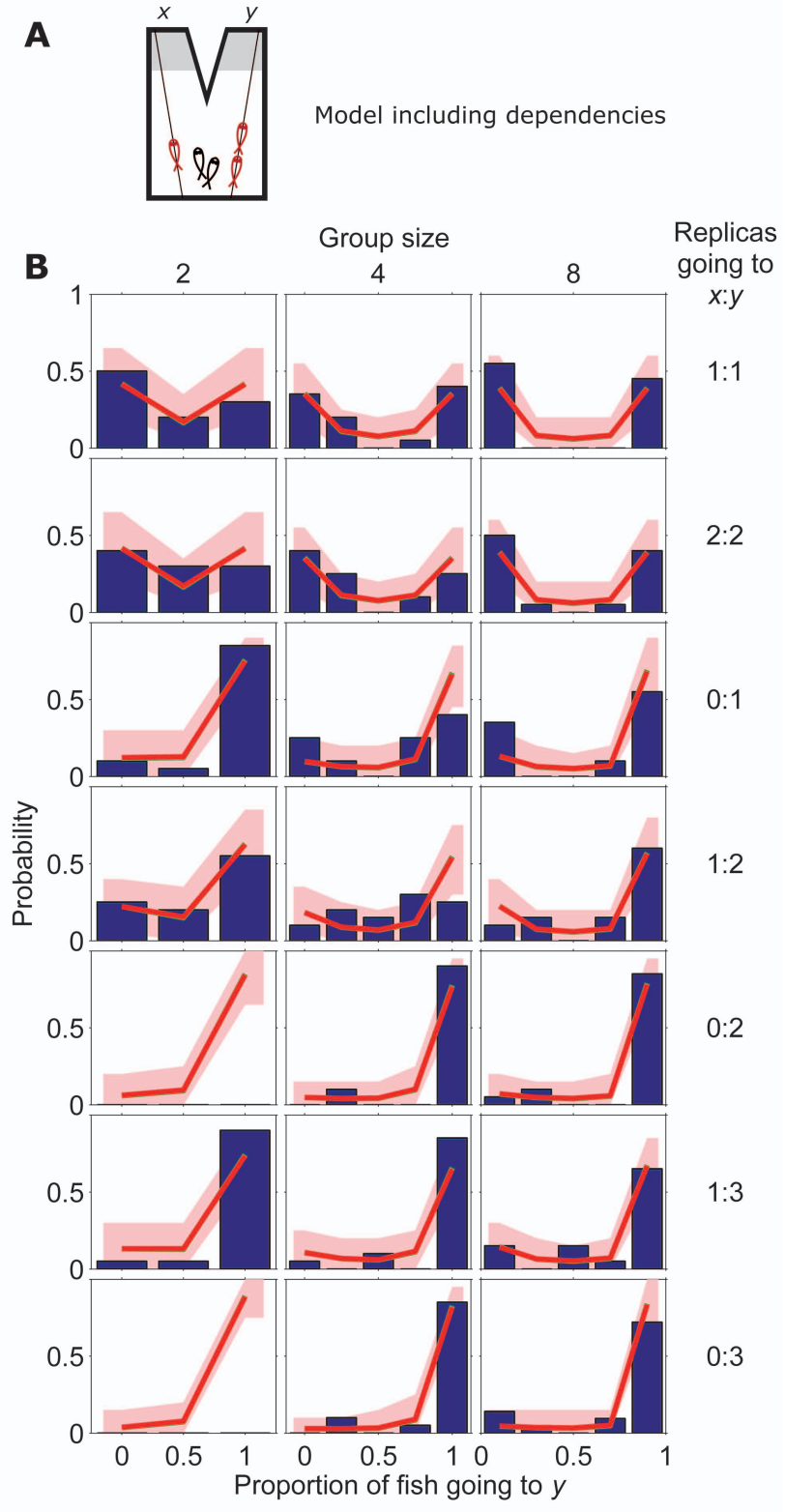


Figure S4: *Comparison between model including dependencies and stickleback choices in symmetric set-up.* (*A*) Schematic diagram of symmetric set-up with a group of sticklebacks (in black) choosing between two identical refugia and with different numbers of replica fish (in red) going to x and y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from the model that takes into account dependencies (red line using $a'_X = 4.9$; red region: 95% confidence interval; green line with $a'_X = 5$). Different graphs correspond to different stickleback group sizes and different number of replicas going to x and y.

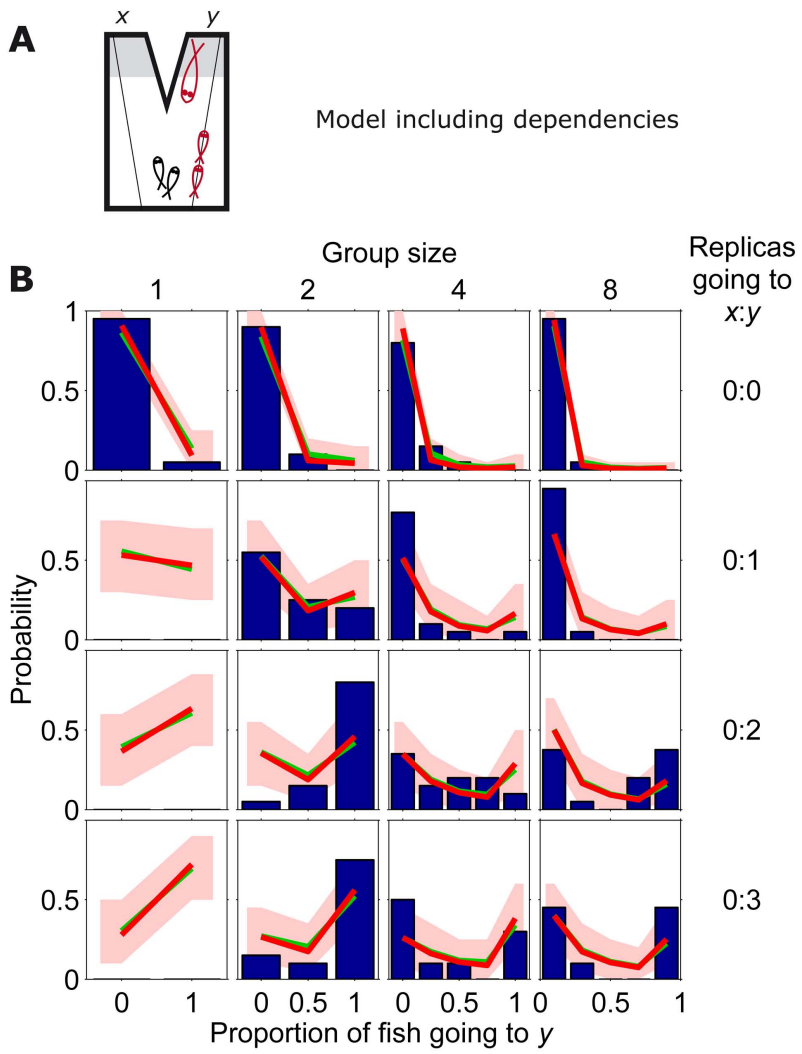


Figure S5: *Comparison between model including dependencies and stickleback choices in asymmetric set-up.* *A*) Schematic diagram of asymmetric set-up (predator at y, large fish depicted in red) with a group of sticklebacks (in black) choosing between two refugia, and replica fish (small fish depicted in red) going to y. (*B*) Experimentally measured statistics of final configurations of fish choices from 20 experimental repetitions [42] (blue histogram) and results from the model that takes into account the dependencies (red line using $a'_X = 8.7$, $a = 9.9$; red region: 95% confidence interval. Green line using $a'_X = 5$ and $a = 6.28$). Different graphs correspond to different stickleback group sizes and different number of replicas going to y.

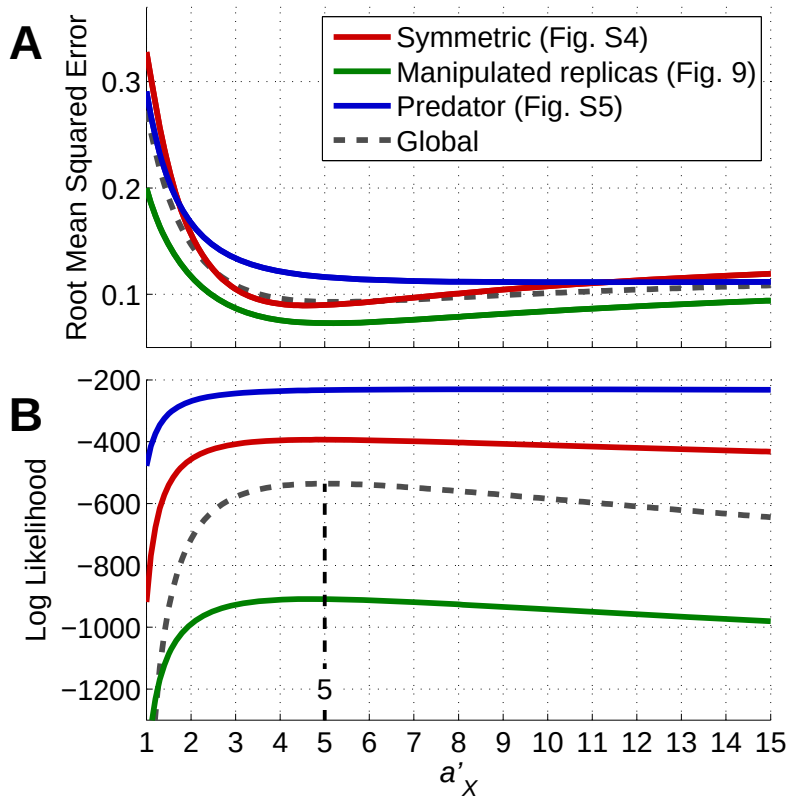


Figure S6: *Goodness of fit of the model including dependencies for different values of* a'_X. **Red:** Symmetric case (data in Fig. S4). **Green:** Case with different replicas at each side (data in Fig. 9. The parameters $a'_{X,\text{replicas}}$ are re-optimized for each value of a'_X). **Blue:** Asymmetric set-up with predator on one side (data in Fig. S5; Parameter a is re-optimized for each value of a'_X). (*A*) Root mean squared error between the data and the probabilities predicted by the model. Grey dashed line shows the mean RMSE for the three cases. The absolute values for each case depend on the shape of the data and are not comparable, only the trends and the position of the minima should be compared. (*B*) Logarithm of the probability that the data come from the model. The height of each curve depends on the number of data for each experiment, only the trend and the position of the maxima should be compared. Grey dashed line shows the sum of the three coloured lines, but shifted by 1000 so that it fits on the scale. The peak of this global probability indicates the value of a'_X that best fits the three datasets ($a'_X = 5$).

Supporting Text

Animal collective behavior from Bayesian estimation and probability matching

Alfonso Pérez-Escudero[1,2,*], Gonzalo G. de Polavieja[1,2,*]

1 Instituto Cajal (CSIC) Madrid, Spain
2 Universidad Autónoma de Madrid, Spain
∗ E-mail: alfonso.perez.escudero@cajal.csic.es, gonzalo.polavieja@cajal.csic.es

Abstract

We present a derivation of the model for the more general case of *M* different options (instead of the 2 options used in the main text). We also discuss some particular cases that give simple expressions while still widely applicable.

Model for *M* options

Let M be the number of possible options, y_m, $m = 1 \dots M$. Each individual estimates the probability that each option is the best one, using its non-social information (C) and the behavior of the other individuals (B). So for one given option, say y_μ, we want to compute

$$P(Y_\mu|C,B), \tag{S1}$$

where Y_μ stands for 'y_μ is the best option'. We can compute the probability in Eq. **S1** using Bayes' theorem,

$$P(Y_\mu|C,B) = \frac{P(B|Y_\mu,C)P(Y_\mu|C)}{\sum_{m=1}^{M} P(B|Y_m,C)P(Y_m|C)}. \tag{S2}$$

Dividing numerator and denominator by the numerator, we get

$$P(Y_\mu|C,B) = \frac{1}{\sum_{m=1}^{M} a_{m\mu} S_{m\mu}}, \tag{S3}$$

where

$$a_{m\mu} = \frac{P(Y_m|C)}{P(Y_\mu|C)} \tag{S4}$$

contains only non-social information, and

$$S_{m\mu} = \frac{P(B|Y_m,C)}{P(B|Y_\mu,C)} \tag{S5}$$

contains the social information. Note that each term of the summation preserves the multiplicative relation between social and non-social information that was also apparent in Eq. 3 of the main text. There may be $M-1$ independent non-social parameters $a_{m\mu}$ in the case that no two options have equal non-social information. But usually this will not be the case, and the number of independent non-social parameters will be lower.

Now we assume independence among behaviors (Eq. 6 in main text), and group all possible behaviors in L classes, $\{\beta_k\}_{k=1}^{L}$ (Eq. 7 in main text). These two assumptions transform Eq. **S5** into

$$S_{m\mu} = \prod_{k=1}^{L} s_{k,m\mu}^{n_k}, \tag{S6}$$

where n_k is the number of individuals performing behavior β_k, and

$$s_{k,m\mu} = \frac{P(\beta_k|Y_m,C)}{P(\beta_k|Y_\mu,C)} \tag{S7}$$

are the reliability parameters for behavior β_k with respect to options y_m and y_μ. There may be up to $L(M-1)$ independent reliability parameters but usually they will not be all independent.

In summary, from Equations **S3** and **S6** we have that

$$P(Y_\mu|C,B) = \left(\sum_{m=1}^{M} a_{m\mu} \prod_{k=1}^{L} s_{k,m\mu}^{n_k} \right)^{-1}. \tag{S8}$$

This equation summarizes the general model applicable to any kind of experiment. In the following sections we consider two particular cases with a much simpler expression.

One basic reliability parameter

The general model in Eq. **S8** depends in general on $L(M-1)$ independent reliability parameters $s_{k,m\mu}$. Here we derive the model for a particular case in which there is only one reliability parameter, s.

First, we consider classes of behaviors (from now on we call them just 'behaviors') that simply consist of choosing a given option. If for example the options are different places, behaviors would be going to each of those places. Therefore, the number of possible behaviors is the same as the number of options, $L = M$. We use the convention that β_j is 'choosing option y_j'. Note that when a behavior is not informative (i.e. its reliability parameter is 1) it has no impact on the model in Eq. **S8**. Therefore, considering this set of behaviors is equivalent to assuming that all other behaviors have reliability parameter equal to 1.

We further assume that $P(\beta_k|Y_m,C)$ only depends on whether $k = m$ or $k \neq m$, so that

$$\begin{aligned} &P(\beta_k|Y_k,C) = P(\beta_l|Y_l,C) \\ &P(\beta_k|Y_m,C) = P(\beta_l|Y_p,C), \quad k \neq m,\ l \neq p \end{aligned} \tag{S9}$$

Note that $P(\beta_k|Y_k, C)$ is the probability that another individual makes the correct choice, and $P(\beta_k|Y_m, C) with\ k \neq m$ is the probability that it makes a wrong choice. So this assumption means that the probability of making the correct choice is the same regardless of which option is actually the correct one. In the case of symmetric choices, in which non-social information C is the same for all options, this relation will hold automatically, not being an extra assumption. It is likely that it also holds for many asymmetric choices. For example, the results for the asymmetric set-up presented in the main text suggest that it holds in that case. We define

$$\begin{aligned} p_c &\equiv P(\beta_k|Y_k, C) \\ p_f &\equiv P(\beta_k|Y_m, C), \quad k \neq m. \end{aligned} \tag{S10}$$

As it only matters whether the behavior matches the correct choice or not, there are only four distinct types of reliability parameters $s_{k,m\mu}$ (Eq. **S7**):

$$\begin{aligned} s_{k,kk} &= \frac{P(\beta_k|Y_k, C)}{P(\beta_k|Y_k, C)} = \frac{p_c}{p_c} = 1 \\ s_{k,ml} &= \frac{P(\beta_k|Y_m, C)}{P(\beta_k|Y_l, C)} = \frac{p_f}{p_f} = 1, \quad k \neq m,\ k \neq l \\ s_{k,km} &= \frac{P(\beta_k|Y_k, C)}{P(\beta_k|Y_m, C)} = \frac{p_c}{p_f} = s, \quad k \neq m \\ s_{k,mk} &= \frac{P(\beta_k|Y_m, C)}{P(\beta_k|Y_k, C)} = \frac{p_f}{p_c} = \frac{1}{s}, \quad k \neq m, \end{aligned} \tag{S11}$$

where

$$s \equiv \frac{p_c}{p_f} \tag{S12}$$

is the basic reliability parameter, equal to the probability that another individual makes the correct choice over the probability that it makes a mistake, for any behavior and for any individual. We regroup the terms inEq. **S8** so that it reflects the different types of $s_{k,m\mu}$ (Eq. **S11**), and get

$$P(Y_\mu|C, B) = \left(\sum_{m=1}^{M} a_{m\mu} s_{m,m\mu}^{n_m} s_{\mu,m\mu}^{n_\mu} \prod_{\substack{k=1 \\ k\neq m \\ k\neq\mu}}^{L} s_{k,m\mu}^{n_k} \right)^{-1}. \tag{S13}$$

Using the relations in Eq. **S11** we have that

$$P(Y_\mu|C, B) = \left(\sum_{m=1}^{M} a_{m\mu} s^{-(n_\mu - n_m)} \right)^{-1}. \tag{S14}$$

Note that the term $m = \mu$ is always equal to 1, so Eq. **S14** is identical to

$$P(Y_\mu|C, B) = \left(1 + \sum_{\substack{m=1 \\ m\neq\mu}}^{M} a_{m\mu} s^{-(n_\mu - n_m)} \right)^{-1}, \tag{S15}$$

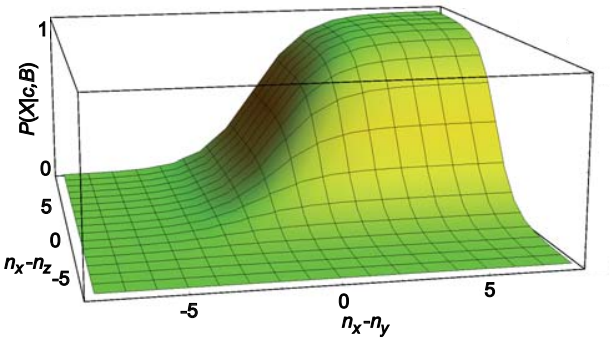


Figure 1. Probability of choosing one of the options for the 3-choice symmetric case.

that has the same structure as the equations presented in the main text.

Symmetric case

In the special case that all options are indistinguishable using non-social information alone (symmetric case), all non-social parameters $a_{m\mu}$ are equal 1 and Eq. **S15** becomes

$$P(Y_\mu|C, B) = \left(1 + \sum_{\substack{m=1 \\ m\neq\mu}}^{M} s^{-(n_\mu - n_m)} \right)^{-1}. \tag{S16}$$

We recall that in this case Eq. **S9** holds automatically, not being an extra assumption.

In the particular case of 3 options, x, y, z, we have

$$P(X|C, B) = \left(1 + s^{-(n_x - n_y)} + s^{-(n_x - n_z)} \right)^{-1}, \tag{S17}$$

and the corresponding expressions for $P(Y|C, B)$ and $P(Z|C, B)$. Fig. 1 shows $P(X|C, B)$ in terms of its two effective variables, $n_x - n_y$ and $n_x - n_z$ (Eq. **S17**).

5
A common rule for three species

José María Nieto - Publicado en El Diario de Valladolid (El Mundo) el 30/10/2010

A common rule for decision making in animal collectives across species

Sara Arganda[1,2], Alfonso Pérez-Escudero[2], and Gonzalo G. de Polavieja[3]

Instituto Cajal, Consejo Superior de Investigaciones Científicas, 28002 Madrid, Spain

Edited by Simon A. Levin, Princeton University, Princeton, NJ, and approved November 2, 2012 (received for review June 21, 2012)

A diversity of decision-making systems has been observed in animal collectives. In some species, choices depend on the differences of the numbers of animals that have chosen each of the available options, whereas in other species on the relative differences (a behavior known as Weber's law), or follow more complex rules. We here show that this diversity of decision systems corresponds to a single rule of decision making in collectives. We first obtained a decision rule based on Bayesian estimation that uses the information provided by the behaviors of the other individuals to improve the estimation of the structure of the world. We then tested this rule in decision experiments using zebrafish (*Danio rerio*), and in existing rich datasets of argentine ants (*Linepithema humile*) and sticklebacks (*Gasterosteus aculeatus*), showing that a unified model across species can quantitatively explain the diversity of decision systems. Further, these results show that the different counting systems used by animals, including humans, can emerge from the common principle of using social information to make good decisions.

collective behavior | public information | probability matching

Sensory data always has some degree of ambiguity, so animals need to make decisions by estimating the properties of the environment from uncertain sensory data (1–5). This estimation has been shown to be close to optimal in many cases, making optimal Bayesian decision making a successful framework shared by behavioral, neurobiological, and psychological studies (1–7).

A richer scenario for decision making takes place when animals move in groups. In this case, the behaviors of other animals are an extra source of information (6–34). Animals of different species have been observed to incorporate this extra information in their decisions in different ways. Some species make decisions that can be explained using the differences of the numbers of animals taking each option (21, 22), others according to the relative differences (Weber's law) (23, 24) or using other rules (25–34). This diversity of decision schemes has translated into a diversity of models (21, 22, 24–34).

To search for a unified framework having the diversity of decision-making schemes as particular cases, we generalized Bayesian decision making to the case of animal collectives. Our previous attempt at building such a theory predicted that the only relevant social information is the difference of the numbers of individuals already choosing each available option, and not the numbers themselves or the relative differences (or Weber's law) (22). However, this theory was limited to the particular case in which only one of the options could be a good option (22). We have now generalized the theory, allowing all available options to be good or bad options. We found that this generalization explains the diversity of decision rules observed in collectives, maintaining the same conceptual and mathematical simplicity, and containing our previous theory as a particular case. We have tested the theory experimentally in decision experiments using zebrafish (*Danio rerio*), but to cover the diversity of decision systems, we have also tested it using rich datasets of decision making in argentine ants (*Linepithema humile*) (24) and three-spined sticklebacks (*Gasterosteus aculeatus*) (25, 26). We found a quantitative match between the theory and the different decision systems of these representative species.

Results

We studied how the behaviors of others should be taken into account to improve the estimations of the structure of the world and make decisions in animal collectives. For a situation with two identical options to choose from (Fig. 1*A*), we looked for the probability that one option, say x, is a good option given that n_x and n_y animals have already chosen options x and y, respectively. We used Bayesian theory to find an approximated analytic expression for this probability as (*SI Text*)

$$P(x \text{ is good}) = \frac{1}{1+as^{-(n_x-kn_y)}}. \quad [1]$$

Parameter a measures the quality of nonsocial information available to the deciding individual, and s measures how reliably an individual that has chosen x indicates to the deciding individual that x is a good option. According to Eq. **1**, the higher the number of individuals that chose option x, n_x, the higher the probability that option x is good for the deciding individual, and more so the higher the reliability s of the information from the individuals that already chose x. However, each individual that chooses y decreases the probability that x is a good option. Parameter k measures the relative impact of these two opposing effects. Individuals need to decide based on the estimated probabilities in Eq. **1**. A common decision rule in animals, from insects to humans, is probability matching, according to which the probability of choosing a behavior is proportional to the estimated probability (35–44),

$$P_x = \frac{P(x \text{ is good})}{P(x \text{ is good}) + P(y \text{ is good})}. \quad [2]$$

This rule is known to be optimal when there is competition for resources (39, 40) and when the estimated probabilities change in time (41–44). Probability matching in Eq. **2**, together with the estimation in Eq. **1**, gives that the probability of choosing x is

$$P_x = \left(1+\frac{1+as^{-(n_x-k\,n_y)}}{1+as^{-(n_y-k\,n_x)}}\right)^{-1}, \quad [3]$$

and $P_y = 1-P_x$ is the probability of choosing y. The main implications of Eq. **3** are apparent in its plot (Fig. 1*B*). First, decision making in collectives is predicted to be different for low and high numbers of individuals. For low numbers, there is a fast

Author contributions: S.A., A.P.-E., and G.G.d.P. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]Present address: Centre de Recherches sur la Cognition Animale, Centre National de la Recherche Scientifique Unité Mixte de Recherche 5169, Université Paul Sabatier, 31062, Toulouse, France.

[2]S.A. and A.P-E contributed equally to this work.

[3]To whom correspondence should be addressed. E-mail: gonzalo.polavieja@cajal.csic.es.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1210664109/-/DCSupplemental.

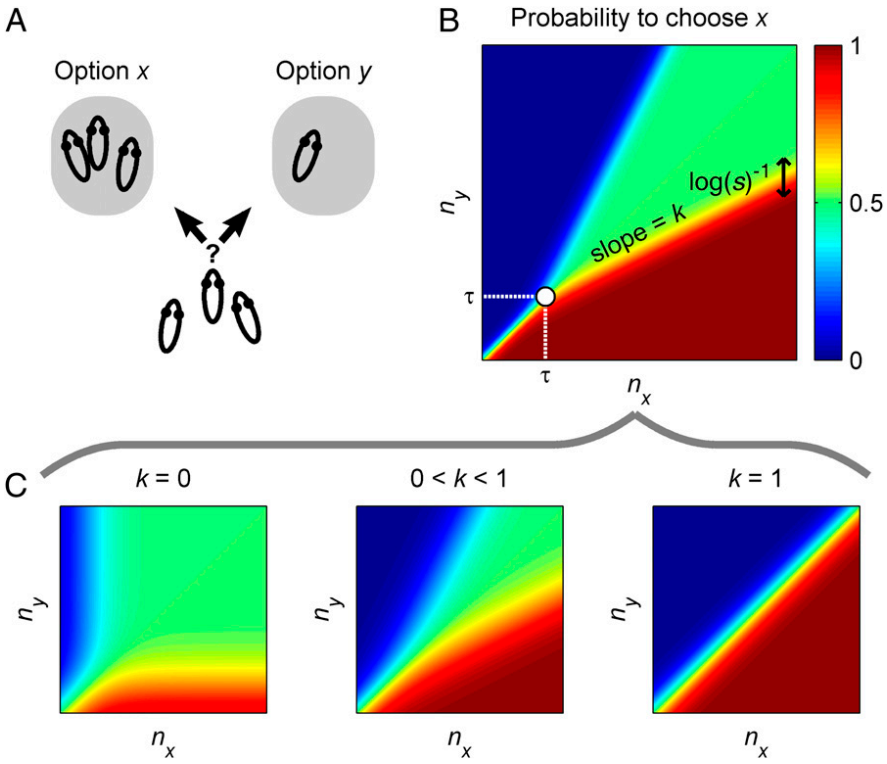


Fig. 1. A general decision-making rule in animal collectives. (*A*) Decision making between two sites when n_x and n_y animals have already chosen sites *x* and *y*, respectively. (*B*) The probability of choosing *x* in the general rule (Eq. **3**), plotted as a function of the animals that have already chosen between the two sites, n_x and n_y. The theory predicts very different structure in the probability for the case of low and high numbers of animals, separated by point $\tau = \log(a)/(\log(s)\ (1-k))$. The rate of change of P_x in the transition regions depends on the reliability parameter *s*, with the width of these regions proportional to $1/\log(s)$. (*C*) Same as *B* but for three different values of parameter *k*: $k = 0$ (*Left*), $0 < k < 1$ (*Center*), and $k = 1$ (*Right*).

transition between preferring one side over the other, whereas for high numbers the transition has an intermediate region with no preference in which the probability has a plateau of value of one-half. There is a clear separation between the low- and high-numbers regimes at the point $\tau = \log(a)/(\log(s)\ (1-k))$ in which the plateau starts (Fig. 1*B*; *SI Text*). Second, in the high-numbers regime, the isoprobability curves are straight lines of slope *k*. We can use this slope to classify three very different scenarios we found to correspond to different experimental datasets: $k = 0$, $0 < k < 1$, and $k = 1$ (Fig. 1*C*).

For $k = 0$, the animals at one option do not impact negatively on the estimated quality of the other option; this can take place, for example, when animals at one option do not seem to have information about the other option. An important prediction for this case is that for high numbers of animals there is a large plateau of probability one-half of choosing each of the two options (Fig. 1*C*, *Left*). To have a significantly higher probability of choosing one option, say *x*, it is then needed not only that $n_x > n_y$ but also to be outside of the large plateau, which means that very few animals have chosen the other option *y*, $n_y < \tau$. A second prediction is that there is a finite number of animals that need to be distinguished; to see this, consider that the probability that option *x* is a good one (Eq. **1**) for $k = 0$ increases monotonically with n_x and converges to 1. The number of animals n_x needed to reach a high probability of 0.95 is given by $\alpha = (\log(a) + \log(1/0.95 - 1))/\log(s)$ (Fig. S1). Beyond α the probability changes very little, thus in practice it is not necessary to count beyond that number. For a wide range of parameters *a* and *s*, α has low values, corresponding to counting up to a low number of animals (Fig. S1).

We have found that wild-type zebrafish (*D. rerio*) in a two-choice setup used for tests of sociability (45, 46) make choices that quantitatively correspond to the predictions of the $k = 0$ case. The setup has three chambers separated by transparent walls; a central chamber with the zebrafish we monitor, and two lateral chambers with different numbers of zebrafish acting as social stimuli (Fig. 2*A*; *Materials and Methods*). An interesting feature of this setup is that it measures the behavior of a single individual when presented with social stimuli, allowing a direct test of the individual decision rule in Eq. **3**. Specifically, we measured the probability that the focal fish chooses each of the two options for a range of configurations (Fig. 2*B*; each dot is the mean of typically $n = 15$ animals). We found that these experimental results correspond to Eq. **3** for $a = 11.2$, $s = 5$, and $k = 0$ (Fig. 2*B*, blue surface) with a robust fit (Fig. S2). To make a more quantitative comparison between theory and experiment, we highlighted several lines on the theoretical surface, using different colors to indicate different numbers of fish at option *y*. Fig. 2*C* compares the probability values for these five lines with the experimental data, showing a close match. The model offers both a quantitative fit to data and a simple explanation of the experimental result. Fish do not choose directly according to the number of other fish, but to how these numbers indicate that a place is a good option, giving a rule of "counting up to 3."

The close match between experimental data and the decision-making model supports that zebrafish behavior corresponds to probabilistic estimations about the quality of sites using social information. However, the processing steps made by the fish brain need not have a one-to-one correspondence with the computational steps in the theory. Instead, a likely option is that zebrafish use simple behavioral rules that approximate good estimations. We found mechanistic models with simple probabilistic attraction rules for individual fish that approximate well the decision-making model and the data (Figs. S3 and S4).

The second case we consider has parameter *k* in the range from 0 to 1. For this range, the estimation that *x* is a good option increases with how many animals have already chosen *x* and decreases,

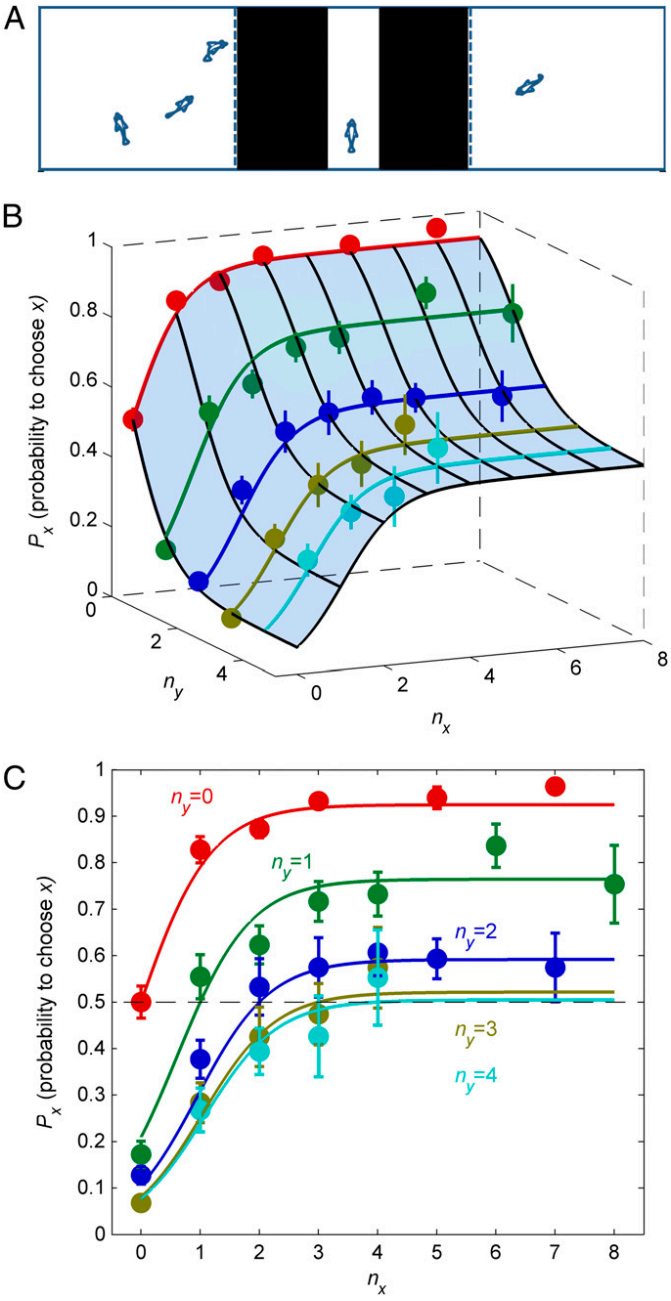


Fig. 2. Zebrafish choices correspond to the general rule of decisions in collectives. (*A*) Focal fish choosing between two sites with different number of zebrafish, separated from the focal fish by glass. (*B*) Probability of choosing option *x* for different numbers of zebrafish at sites *x* and *y*, n_x and n_y. Theoretical probabilities for $a = 11.2$ and $s = 5$, and $k = 0$ in Eq. **3** represented as a surface and experimental data represented as dots indicating the mean value of typically 15 animals at each configuration. Different dot colors correspond to different values of n_y and bars are SEM. (*C*) Same as *B* but plotted only as a function of n_x and different colors representing the value of n_y.

although at a slower rate, with how many have chosen option *y*. This situation might be common, for example, in food search. Animals choosing one option can indicate that there is a food source in that direction, but also that there might not be a food source at the other option. In this case, the probability of choosing *x* has a plateau in which both options are equally likely, but increasing the number of animals that have chosen *x*, n_x, reaches a transition region of rapid increase in probability (Fig. 1*B*). This transition region follows a straight line of slope *k* in the probability plot (Fig. 1*B*). This line obeys for high number of animals that $n_y \approx k\, n_x$. This is a Weber law (23, 24), according to which the just-noticeable difference between two groups is proportional to the total number of individuals. Indeed, if we substitute $n_y \approx k\, n_x$ into $\Delta N/N \equiv (n_x - n_y)/(n_x + n_y)$ we obtain a constant of value $(k-1)/(k+1)$. A second prediction of the model is that decisions should deviate from Weber behavior at low numbers (below the transition point τ in Fig. 1*B*).

We have found that decisions made by the Argentine ant (*L. humile*) correspond to the case $0 < k < 1$. Ants' choices to turn left and right have been recorded by Perna et al. (24), and we found that they have choice probabilities well described by Eq. **3**, except that experimental probabilities do not reach values as close to 0 or 1 as the theory. This difference might be due simply to the fact that ants are not always making turn decisions based on pheromones, but responding to other factors, such as roughness of terrain or collisions with other ants. We therefore considered that ants choose at random with a given probability and otherwise make a decision according to Eq. **3** (Eq. **4**). This modification only introduces an overall rescaling in the probabilities, so all structural features described below are present in Eq. **3** (Fig. S5). We obtain a good correspondence with data for high (Fig. 3*A*) and low numbers of animals (Fig. 3*B*) with a fit that is robust (Fig. S6). The experimental data are smoother than the theory, without a central plateau, but still with a close correspondence, as also shown in the following analysis. According to Weber's law, isoprobability curves should be horizontal lines in the $\Delta N/N \equiv (n_x - n_y)/(n_x + n_y)$ vs. $N \equiv n_x + n_y$ plane, and this is true both for the theory and experiments for high numbers of total animals *N* (Fig. 3*C*). The advantage of this plot is that it magnifies the region of low *N*, where the data deviate from Weber's law similarly to the theoretical prediction. A further quantitative analysis revealing the close correspondence between theory and data are shown in Fig. 3*D*. We performed a linear fit to the experimental probability along the lines of constant $n_x + n_y$ depicted in Fig. 3*D Inset*. The slope of each linear fit was then plotted against the total number of animals *N* (Fig. 3*D*, blue dots). The experimental data has a very close correspondence with the theoretical values in this plot (Fig. 3*D*, red line). For a high number of animals, both theory and data show Weber behavior, corresponding in this logarithmic plot to a straight line with slope −1 (Fig. 3*D*, black line) (24). Interestingly, for low numbers of animals, the theoretical prediction of a deviation from Weber behavior corresponds to the data.

The last case we consider has $k = 1$, for which Eq. **3** depends only on the variable $\Delta N \equiv n_x - n_y$. This situation could take place when there is a high probability that only one of the options is good, and those animals choosing *x* indicate that *x* may be the good one in a similar way that those choosing *y* may indicate that *x* might not be the good one. We have previously shown (22) that the simple decision rule $P_x = 1/(1 + a\, s^{-\Delta N})$ explains well a large dataset of collective decisions in sticklebacks, *G. aculeatus* (25, 26). In these experiments, animal groups were made to choose in two-choice setups with different combinations of social and nonsocial information (Fig. 4*A*, *Far Left*). Interestingly, Eq. **3** has the simple rule $P_x = 1/(1 + a\, s^{-\Delta N})$ as a particular case for $k = 1$ (*SI Text*). Indeed, all experimental results (blue histograms in Fig. 4*A* and Fig. S7) are fit using Eq. **3** with parameters $s = 2.5$, $k = 1$ (Fig. 4*A*, red lines). Additionally, for low numbers of animals (up to τ in Fig. 1*B*), an approximated ΔN rule can also be found for any value of *k* but with different values of the nonsocial reliability parameter *a* (*SI Text*). Therefore, the stickleback data can be fit with any value of *k* (green and blue lines in Fig. 4*A* and Fig. S7 for $k = 0.5$ and $k = 0$, respectively), with robust fits (Fig. S8). The reason why in this case *k* can have any value is that its main effect is to control the slope of the boundaries of the plateau of probability 0.5, which is not present in the experimentally explored region of the stickleback dataset (Fig. 4*B*, white triangle). Still, all these fits

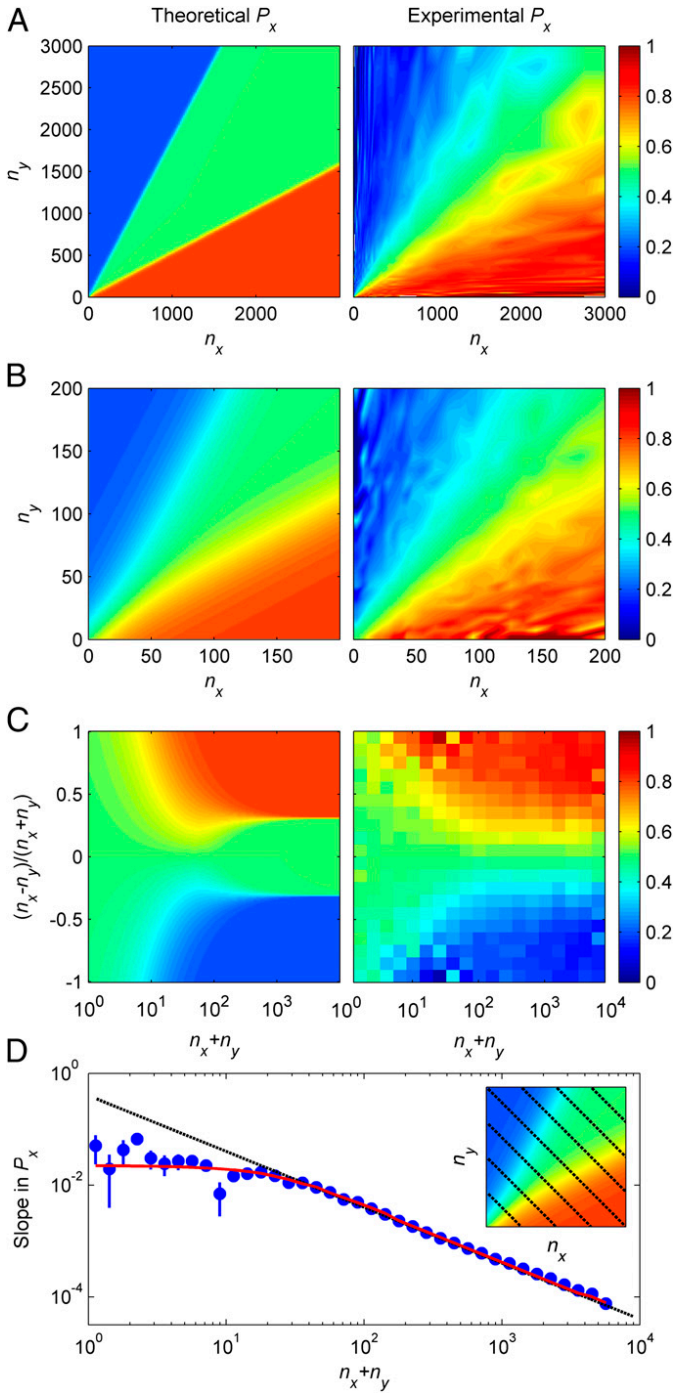


Fig. 3. Ant choices correspond to the general rule of decisions in collectives. (*A*) Probability of choosing option x as a function of how many ants have previously been at locations x and y, n_x and n_y, for theory (*Left*) using Eq. **4** with $a = 2.5$, $s = 1.07$, $k = 0.53$, $p_{rand} = 0.39$, and experiments (*Right*) from Perna et al. (24). (*B*) Detail of *A*. (*C*) Same as *A* but represented as a function of $\Delta N/N$ and N. (*D*) Slope of the probability of choosing x in *A* as obtained from a linear fit along the lines depicted in *Inset*. Experimental values (blue dots; error bars are 95% confidence interval), theory (red line), and Weber's law (black line).

have in common an effective ΔN rule for the experimental region (Fig. 4*B*), giving strong support to this rule in this dataset.

Discussion

Our results support that estimation by the brain using social information to counteract the ambiguity of sensory data is a fundamental principle in collective decision making. The theory explains also the diversity in number discrimination schemes used in collective decisions, including counting up to a given number of animals, counting the difference of animals choosing among options, ΔN, or the relative difference, $\Delta N/N$, as well as observed deviations from these ideal cases and the existence of different counting regimes for high and low numbers as observed in many species, including humans (47, 48). A single mathematical rule contains all these cases and can be used as a first-principles approach to quantitatively study decisions in animal collectives.

One important ingredient of our theory is the use of probability matching (Eq. **2**); for symmetric decisions, it implies a functional form of the type $P_x = f(x,y)/(f(x,y)+f(y,x))$. Our model in Eq. **3** is a particular case of this function, with $f(x,y)$ derived from an approximation to Bayesian estimation. Interestingly, many previous approaches derive from the form $P_x = f(x)/(f(x)+f(y))$ (21, 22, 27, 28), which is also a particular case of $P_x = f(x,y)/(f(x,y)+f(y,x))$, and therefore compatible with probability matching. In other cases, the basic form $P_x = f(x)/(f(x)+f(y))$ has been modified by adding constant terms (29, 30) or an extra function (25), as $P_x = f(x)/(f(x)+f(k))$, with k a constant when animals have access to a single choice (31, 32, 34). Weber behavior can also be seen as a particular case. It has been previously described using a function (24) that can be expressed as $f(x,y) = 1/2 + \delta(n_x - n_y)/(n_x + n_y)$, with δ between 0 and 1/2. This function obeys $f(x,y)+f(y,x) = 1$, so in this case $P_x = f(x,y)$, following Weber behavior.

These previous functions are very useful when applied to particular datasets because they may use few parameters in these conditions. In particular, our previous model (22), a particular case of Eq. **3** (*SI Text*), used only one parameter in the symmetric experiments with sticklebacks, and a model with two parameters described the ants dataset (24). However, these two models cannot fit the three datasets or even two of them (Fig. S9 *A* and *D* and S10). For the zebrafish data in Fig. 2, none of the previously proposed functions (21, 22, 24, 27–29) give a good fit of the plateau in the data (Fig. S9). Our approach has been developed to be applied in very different species and conditions, here tested for three large datasets in three different species. One important factor in this ability to describe different datasets is that our basic function $f(x,y)$ has a term $s^{-(n_x - k\, n_y)}$ that captures how the estimated quality of an option depends not only on the animals choosing that option but also on the animals choosing the other option. These two sources of information are balanced by parameter k, and different datasets are found to correspond to different balances k.

Previous functions describing ant foraging include a constant term that represents a threshold of pheromone concentration below which ants do not react (24, 27, 28). In this way, these functions can describe the deviation from Weber's law at low pheromone concentration (24). In our case, the theory naturally shows this behavior as one more particular case of the predicted difference between a low and high number of animals. Comparing the two approaches, it is interesting to consider that the behavior for low numbers that is predicted from estimation theory can be achieved in ants using a threshold of pheromone concentration.

An advantage of our approach is that the form of the function f is derived for any type of setup simply from estimation given nonsocial sensory data and the behaviors of others (*SI Text*). For example, we predict for a symmetric setup with N options a generalization of Eq. **1** of the form

$$P(x \text{ is good}) = \frac{1}{1 + a s^{-(n_x - k\, M)}},$$

with $M = \sum_{i \neq x}^{N} n_i$ the total number of animals choosing any option except x (see Eq. **S10** for the more general case of asymmetric choices).

A further advantage is that the parameters a, s, and k are not only fitting parameters but have expressions, Eqs. **S4**, **S9**, and

ECOLOGY

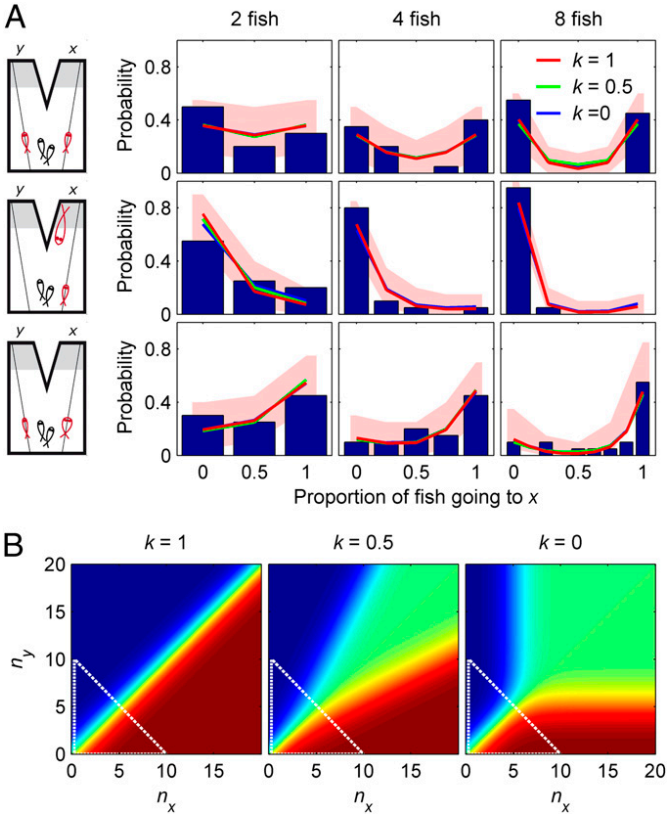


Fig. 4. Stickleback choices correspond to the general rule of decisions in collectives. (*A*) Probability of finding a final proportion of sticklebacks choosing option *x* (blue histograms are experimental results from refs. 25 and 26 and theoretical values as lines for $k = 1$, $k = 0.5$, and $k = 0$) for different group sizes (two, four, and eight fish) and for three types of setups: a symmetric setup with different numbers of replica fish going to *x* and *y* (*Top*), a setup with a replica predator at *x* and different replica fish going to *x* (*Middle*), and a symmetric setup with modified replica fish (*Bottom*). See model parameters and 68 additional experiments with fits in Fig. S7. (*B*) Theoretical P_x for $k = 1$, $a = 1$ (*Left*), $k = 0.5$, $a = 5$ (*Center*), and $k = 0$, $a = 224$ (*Right*), and $s = 2.5$ in the three cases. All models require an effective ΔN rule to compare with data for the number of animals used in experiments (triangle).

S16, respectively, which give additional predictions. For example, the social reliability parameter is given by

$$s = \frac{P(\beta|X,C)}{P(\beta|\overline{X},C)},$$

with β a given animal behavior. This expression means that the social reliability parameter s is higher for a behavior β that is produced with high probability when x is a good option, and with very low probability when it is not a good option. Among all behaviors, those with higher s allow an individual to obtain a higher probability that option x is a good one (Eq. **1**), so we expect them to have a larger effect on collective decision making.

Another advantage of an approach based on a theory of estimation is that generalizations of the theoretical expressions can be envisaged deriving models using fewer assumptions. For example, including dependencies in the behaviors of the other individuals and explicit space and time variables should be natural extensions of the theory.

Materials and Methods

Experimental Protocol for Zebrafish. All procedures met with European Economic Community (EEC) guidelines for animal experiments under Directive 86/609/EEC. Experimental procedures were approved by the Bioethics Subcommittee of Consejo Superior de Investigaciones Científicas. We used wild-type adult zebrafish, *D. rerio*, of both sexes. Fish were acclimatized to the setup water for 1 d before the experiments (Fig. S11). At 1 h before the experiment, each fish was isolated and fed to ensure uniform nutritional status across individuals. A focal fish entered the setup and swam freely in a central chamber between two social chambers with different number of fish and separated from the choice chamber by glass. Once a fish had been recorded for 5 min it could be placed in one of the lateral chambers as a social stimulus for another fish. The fish in the lateral chambers were interchanged between trials to ensure uniformity, and sides were randomized. The central chamber of the setup was washed between trials to remove odor traces. We computed the probability P_x as the fraction of time the focal fish spent on the black region close to one of the social chambers, *x*. This fraction of time converges to P_x for a fish that makes repeated decisions choosing *x* (*y*) with probability P_x ($P_y = 1 - P_x$). A total of 238 fish were tested only once. To test the effect of previous experience, another 233 trials were performed with fish that were tested several times. We found no significant difference between the two groups in the mean times spent at each side (Fig. S12), so all data were pooled for Fig. 2.

Model with Noise Added to the Decision Rule. The model in Eq. **3** has a good agreement with data from experiments using the Argentine ant, *L. humile* (24), except that experimental probabilities do not reach values as close to 0 or 1 as the theory. To account for the experimental data, we made a simple modification of the model by assuming that the ant has some probability p_{rand} of making the decision at random motivated by unknown factors. Then, with probability $(1 - p_{\text{rand}})$, the ant makes the decision according to Eq. **3**. Therefore, the probability of turning toward *x* is

$$P_x = \frac{p_{\text{rand}}}{2} + (1 - p_{\text{rand}})\left(1 + \frac{1 + as^{-(n_x - k\,n_y)}}{1 + as^{-(n_y - k\,n_x)}}\right)^{-1}. \quad [4]$$

The parameters that best fit the ant data are $a = 2.5$, $s = 1.07$, $k = 0.53$, and $p_{\text{rand}} = 0.39$. This same model can be applied to the zebrafish and stickleback datasets, but in these cases the best fit is obtained for $p_{\text{rand}} \approx 0$, which corresponds to Eq. **3**.

Analysis of the Ants Dataset. Both the raw dataset and preprocessing routines were provided by Perna et al. (24), and we used their data assuming no evaporation of pheromone (this assumption does not change the results significantly). We calculated from the data the probability of turning right or left, not a continuous angle, to compare directly to our predicted probabilities. To reduce the noise in the experimental maps of Fig. 3, we symmetrized the data so that the probability shown at point (n_x, n_y) is obtained as $(P_x(n_x,n_y) + (1 - P_x(n_y,n_x)))/2$.

Experimental data from Perna et al. (24) measures a quantity that is proportional to the number of ants previously at the left/right of the deciding ant, not directly the numbers, so the number of ants (n_x, n_y) used in the plots are related to the actual number of ants that count for the decision $(n_{x,\text{true}}, n_{y,\text{true}})$ by $n_x = \lambda\, n_{x,\text{true}}$, $n_y = \lambda\, n_{y,\text{true}}$, where λ is an unknown proportionality constant. This relation means that the model still applies but with $s = s_{\text{true}}{}^{\lambda}$, where s_{true} is the actual value of the reliability parameter.

Asymmetric models. For the case of sticklebacks deciding in the presence of a predator (Fig. 4A, *Middle*), we derived the more general asymmetric version of Eq. **1**, Eq. **S22** (Figs. S13 and S14). We also derived Eq. **S23** for the case with two different types of replicas (Fig. 4A, *Bottom*).

Fitting Procedures. To fit the model's parameters to the data, we performed 2D exhaustive searches in the space of parameters. For functions with more than two parameters, we performed the search successively with all possible pairs of parameters. In these cases we repeated the fit several times starting from different initial conditions, always getting the same final result.

ACKNOWLEDGMENTS. We thank Andrea Perna for sharing the raw ant data before publication; David Sumpter, Jacques Gautrais, Raphael Jeanson, Matthieu Moreau, and members of G.G.d.P.'s laboratory for fruitful discussions; and two reviewers for constructive comments. Funding for this work was provided by the Ministerio de Ciencia e Innovacion (MICINN, Spain) as Plan Nacional and as partners of the ERASysBio+ initiative (G.G.d.P.) supported under the European Union European Research Area Networks (ERA-NET) Plus scheme in Framework Program 7, including contracts to S.A and A.P.E. A.P.E. is the recipient of a Formación de Profesorado Universitario (FPU) fellowship from MICINN (Spain).

1. Knill DC, Pouget A (2004) The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends Neurosci* 27(12):712–719.
2. Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
3. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND (2011) How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022):1279–1285.
4. Biernaskie JM, Walker SC, Gegear RJ (2009) Bumblebees learn to forage like Bayesians. *Am Nat* 174(3):413–423.
5. McNamara JM, Green RF, Olsson O (2006) Bayes theorem and its applications in animal behaviour. *Oikos* 112(2):243–251.
6. Banerjee AV (1992) A simple model of herd behavior. *Q J Econ* 107(3):797–817.
7. Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *J Polit Econ* 100(5):992–1026.
8. Valone TJ, Templeton JJ (2002) Public information for the assessment of quality: A widespread social phenomenon. *Philos Trans R Soc Lond B Biol Sci* 357(1427):1549–1557.
9. Giraldeau LA, Valone TJ, Templeton JJ (2002) Potential disadvantages of using socially acquired information. *Philos Trans R Soc Lond B Biol Sci* 357(1427):1559–1566.
10. Wagner RH, Danchin E (2010) A taxonomy of biological information. *Oikos* 119(2): 203–209.
11. Clark C, Mangel M (1986) The evolutionary advantages of group foraging. *Theor Popul Biol* 30(1):45–75.
12. Coolen I, van Bergen Y, Day RL, Laland KN (2003) Species difference in adaptive use of public information in sticklebacks. *Proc Biol Sci* 270(1531):2413–2419.
13. van Bergen Y, Coolen I, Laland KN (2004) Nine-spined sticklebacks exploit the most reliable source when public and private information conflict. *Proc Biol Sci* 271(1542):957–962.
14. Rieucau G, Giraldeau LA (2009) Persuasive companions can be wrong: The use of misleading social information in nutmeg mannikins. *Behav Ecol* 20(6):1217–1222.
15. Lima SL (1995) Collective detection of predatory attack by social foragers: Fraught with ambiguity? *Anim Behav* 50(4):1097–1108.
16. Nordell SE, Valone TJ (1998) Mate choice copying as public information. *Ecol Lett* 1(2): 74–76.
17. Magurran AE, Higham A (2010) Information transfer across fish shoals under predation threat. *Ethology* 78(2):153–158.
18. Danchin E, Giraldeau LA, Valone TJ, Wagner RH (2004) Public information: From nosy neighbors to cultural evolution. *Science* 305(5683):487–491.
19. Dall SRX, Giraldeau LA, Olsson O, McNamara JM, Stephens DW (2005) Information and its use by animals in evolutionary ecology. *Trends Ecol Evol* 20(4):187–193.
20. Dugatkin LA (1992) Sexual selection and imitation: Females copy the mate choice of others. *Am Nat* 139(6):1384–1389.
21. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc Natl Acad Sci USA* 108(6):2312–2315.
22. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
23. Gómez-Laplaza LM, Gerlai R (2011) Can angelfish (*Pterophyllum scalare*) count? Discrimination between different shoal sizes follows Weber's law. *Anim Cogn* 14(1):1–9.
24. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.
25. Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J (2008) Quorum decision-making facilitates information transfer in fish shoals. *Proc Natl Acad Sci USA* 105(19): 6948–6953.
26. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.
27. Goss S, Aron S, Deneubourg JL, Pastels JM (1989) Self-organized shortcuts in the Argentine ant. *Naturwissenschaften* 76(12):579–581.
28. Deneubourg J, Aron S, Goss S, Pasteels J (1990) The self-organizing exploratory pattern of the argentine ant. *J Insect Behav* 3(2):159–168.
29. Meunier H, Leca JB, Deneubourg JL, Petit O (2006) Group movement decisions in capuchin monkeys: The utility of an experimental study and a mathematical model to explore the relationship between individual and collective behaviours. *Behaviour* 143 (12):1511–1527.
30. Tero A, Kobayashi R, Nakagaki T (2007) A mathematical model for adaptive transport network in path finding by true slime mold. *J Theor Biol* 244(4):553–564.
31. Gallup AC, et al. (2012) Visual attention and the acquisition of information in human crowds. *Proc Natl Acad Sci USA* 109(19):7245–7250.
32. Sumpter DJT, Pratt SC (2009) Quorum responses and consensus decision making. *Philos Trans R Soc Lond B Biol Sci* 364(1518):743–753.
33. Jeanson R, et al. (2005) Self-organized aggregation in cockroaches. *Anim Behav* 69(1): 169–180.
34. Bousquet CAH, Sumpter DJT, Manser MB (2011) Moving calls: A vocal mechanism underlying quorum decisions in cohesive groups. *Proc R Soc B Biol Sci* 278(1711): 1482–1488.
35. Behrend ER, Bitterman ME (1961) Probability-matching in the fish. *Am J Psychol* 74(4): 542–551.
36. Greggers U, Menzel R (1993) Memory dynamics and foraging strategies of honeybees. *Behav Ecol Sociobiol* 32(1):17–29.
37. Kirk KL, Bitterman ME (1965) Probability-learning by the turtle. *Science* 148(3676): 1484–1485.
38. Wozny DR, Beierholm UR, Shams L (2010) Probability matching as a computational strategy used in perception. *PLOS Comput Biol* 6(8):7.
39. Fretwell S, Lucas H (1969) On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor* 19(1):16–36.
40. Houston A, McNamara J (1987) Switching between resources and the ideal free distribution. *Anim Behav* 35(1):301–302.
41. Vulkan N (2000) An economist's perspective on probability matching. *J Econ Surv* 14 (1):101–118.
42. Staddon J (1983) *Adaptive Behavior and Learning* (Cambridge Univ Press, Cambridge, UK).
43. Gaissmaier W, Schooler LJ (2008) The smart potential behind probability matching. *Cognition* 109(3):416–422.
44. Green CS, Benson C, Kersten D, Schrater P (2010) Alterations in choice behavior by manipulations of world model. *Proc Natl Acad Sci USA* 107(37):16401–16406.
45. Engeszer RE, Ryan MJ, Parichy DM (2004) Learned social preference in zebrafish. *Curr Biol* 14(10):881–884.
46. Wright D, Krause J (2006) Repeated measures of shoaling tendency in zebrafish (*Danio rerio*) and other small teleost fishes. *Nat Protoc* 1(4):1828–1831.
47. Feigenson L, Dehaene S, Spelke E (2004) Core systems of number. *Trends Cogn Sci* 8(7):307–314.
48. Dehaene S, Izard V, Spelke E, Pica P (2008) Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science* 320(5880): 1217–1220.

Supporting Information

Arganda et al. 10.1073/pnas.1210664109

SI Text

Here we give the derivation of Eq. **1**, the derivation of a more general equation for an asymmetric setup (Fig. 4*A*, *Middle* row), and the derivation of an equation for a symmetric setup but with different types of animals to follow (Fig. 4*A*, *Bottom* row). Additionally, we show that the model in ref. 1 is a particular case of Eq. **3**. We also present the derivation of an expression for the point τ separating the low-number and high-number decision behaviors, and proof of the approximate ΔN rule for low N.

Derivation of Eq. 1. The following derivation follows similar steps to our derivation in ref. 1, except for the key difference that animals now estimate the probability that different options are good instead of the best. This simple difference makes the theory more general with previous results in ref. 1 only a particular case, as shown in a subsequent section.

Consider a focal individual making a decision among several options (x, y, z, . . .). To make this decision, it estimates the probability that each option is a good choice. "Good" may refer to the presence of food, shelter, absence of predators, or any other feature. To perform this estimation, the individual uses the information of the environment gathered directly by its sensors (nonsocial information, C), and the behaviors of the other individuals (social information, B). The probability that a given option (say, option x) is a good choice, given both nonsocial and social information is

$$P(X|C,B), \qquad \textbf{[S1]}$$

where X stands for "x is a good choice." We can compute this probability using Bayes′ theorem,

$$P(X|C,B) = \frac{P(B|X,C)P(X|C)}{P(B|X,C)\,P(X|C) + P(B|\overline{X},C)\,P(\overline{X}|C)}, \qquad \textbf{[S2]}$$

where $\overline{X}$ stands for "x is not a good choice." Dividing the numerator and denominator of Eq. **S2** by the numerator, we get

$$P(X|C,B) = \frac{1}{1 + a_x S_x}, \qquad \textbf{[S3]}$$

with

$$a_x = \frac{P(\overline{X}|C)}{P(X|C)} \qquad \textbf{[S4]}$$

and

$$S_x = \frac{P(B|\overline{X},C)}{P(B|X,C)}, \qquad \textbf{[S5]}$$

where we use the subindex x to indicate that it refers to the estimation for option x. Each of the options has a set of equations like Eqs. **S3–S5**. Note that a_x only contains nonsocial information (C), so we call it a nonsocial term, whereas the social information (B) is contained in the social term, S_x. A practical version of Eq. **S3** is obtained using the approximation that the focal individual does not take into account the correlations among the rest of individuals (however, see ref. 1 for a treatment of these correlations). This assumption implies that the probability of a given set of behaviors is equal to the product of the probabilities of individual behaviors. We apply this to the probabilities needed to compute S_x in Eq. **S5**,

$$P(B|X,C) = Z\prod_{i=1}^{N} P(b_i|X,C), \qquad \textbf{[S6]}$$

where B is the set of behaviors of the other N animals at the time the focal individual is choosing, $B = \{b_i\}_{i=1}^{N}$, and b_i denotes the behavior of individual i. Z is a combinatorial term counting the number of possible decision sequences leading to the set of behaviors B, that will cancel out below. Substituting Eq. **S6**, and an analogous expression for $P(B|\overline{X},C)$, into Eq. **S5**, we get

$$S_x = \prod_{i=1}^{N} \frac{P(b_i|\overline{X},C)}{P(b_i|X,C)}. \qquad \textbf{[S7]}$$

A more useful expression is obtained if we consider, instead of the full individual behaviors (b_i) with all their details, a set of behavioral classes that group together the behaviors that contain similar information about the choice. For example, in a two-choice setup, useful behavioral classes might be "choosing x" (denoted as β_x) and "choosing y" (β_y). Consider in general L behavioral classes $\{\beta_j\}_{j=1}^{L}$. We do not here consider animals to have individual differences, so all have the same probabilities for each behavior; for example, the same $P(\beta_1|X,C)$ and $P(\beta_1|\overline{X},C)$ for behavior β_1, which means that if the first n_1 individuals are performing behavior β_1, we have $\prod_{i=1}^{n_1}\frac{P(b_i|\overline{X},C)}{P(b_i|X,C)} = \left(\frac{P(\beta_1|\overline{X},C)}{P(\beta_1|X,C)}\right)^{n_1}$. We can then write Eq. **S7** as

$$S_x = \prod_{j=1}^{L} s_{xj}^{-n_j}, \qquad \textbf{[S8]}$$

where n_j is the number of individuals performing behavior β_j, and

$$s_{xj} = \frac{P(\beta_j|X,C)}{P(\beta_j|\overline{X},C)}. \qquad \textbf{[S9]}$$

To summarize, the probability that option x is a good choice is, using Eqs. **S3** and **S8**,

$$P(X|C,B) = \left(1 + a_x\prod_{j=1}^{L} s_{xj}^{-n_j}\right)^{-1}, \qquad \textbf{[S10]}$$

with a_x in Eq. **S4** and s_{xj} in Eq. **S9**.

The zebrafish experiments in the main text were performed in a setup with two identical sites to choose from, except for the number of animals at each site, n_x and n_y. The focal animal can observe two types of behaviors: stay at x (β_x) and stay at y (β_y). Eq. **S10** then reduces to

$$P(X|C,B) = \frac{1}{1 + a_x s_{xx}^{-n_x} s_{xy}^{-n_y}}. \qquad \textbf{[S11]}$$

Similarly, for option y the estimation is

$$P(Y|C,B) = \frac{1}{1 + a_y s_{yy}^{-n_y} s_{yx}^{-n_x}}. \quad \textbf{[S12]}$$

The nonsocial information for the two sites x and y is identical by experimental design, so

$$\begin{aligned} P(X|C) &= P(Y|C) \\ P(\overline{X}|C) &= P(\overline{Y}|C) \end{aligned}. \quad \textbf{[S13]}$$

The relations in Eq. **S13**, together with Eq. **S4**, mean that $a_x = a_y$. For notational simplicity, we then define

$$a \equiv a_x = a_y. \quad \textbf{[S14]}$$

The symmetry of the setup also implies the following relations

$$\begin{aligned} P(\beta_x|X,C) &= P(\beta_y|Y,C) \\ P(\beta_x|\overline{X},C) &= P(\beta_y|\overline{Y},C) \\ P(\beta_x|Y,C) &= P(\beta_y|X,C) \\ P(\beta_x|\overline{Y},C) &= P(\beta_y|\overline{X},C) \end{aligned}. \quad \textbf{[S15]}$$

In an idealized situation in which the only possible behaviors were "stay at x" and "stay at y," we would have that $P(\beta_x|X,C) = 1 - P(\beta_y|X,C)$. Because real behaviors are much more complex, and different behaviors can exist, these two probabilities will not sum 1 in general.

According to Eqs. **S15** and **S9**, we have that $s_{xx} = s_{yy}$ and $s_{xy} = s_{yx}$. It is then useful to define

$$\begin{aligned} s &\equiv s_{xx} = s_{yy} \\ k &\equiv -\frac{\log(s_{xy})}{\log(s_{xx})} = -\frac{\log(s_{yx})}{\log(s_{yy})} \end{aligned}. \quad \textbf{[S16]}$$

Using Eqs. **S14** and **S16**, we can write Eqs. **S11** and **S12** as

$$\begin{aligned} P(X|C,B) &= \frac{1}{1 + a s^{-(n_x - k n_y)}} \\ P(Y|C,B) &= \frac{1}{1 + a s^{-(n_y - k n_x)}} \end{aligned}, \quad \textbf{[S17]}$$

obtaining Eq. **1**. Note that $s = P(\beta_x|X,C)/P(\beta_x|\overline{X},C) = P(\beta_y|Y,C)/P(\beta_y|\overline{Y},C)$, that is, the probability of choosing one option when it is a good choice over the probability of choosing it when it is a bad choice. Therefore, parameter s measures how reliable are the choices of each of the other individuals.

The probability of choosing x or y is then obtained using probability matching, Eq. **2** to give Eq. **3**,

$$P_x = \left(1 + \frac{1 + a s^{-(n_x - k n_y)}}{1 + a s^{-(n_y - k n_x)}}\right)^{-1}. \quad \textbf{[S18]}$$

Derivation of a More General Equation for an Asymmetric Setup. In the case of an asymmetric setup (as in Fig. 4*A*, *Middle* row), the nonsocial information for the two sites x and y is different, so

$$\begin{aligned} P(X|C) &\neq P(Y|C) \\ P(\overline{X}|C) &\neq P(\overline{Y}|C) \end{aligned}. \quad \textbf{[S19]}$$

The relations in Eq. **S19** mean that $a_x \neq a_y$, as is clear from its definition in Eq. **S4**.

In the symmetric case, we used the relations in Eq. **S15**,

$$\begin{aligned} P(\beta_x|X,C) &= P(\beta_y|Y,C) \\ P(\beta_x|\overline{X},C) &= P(\beta_y|\overline{Y},C) \\ P(\beta_x|Y,C) &= P(\beta_y|X,C) \\ P(\beta_x|\overline{Y},C) &= P(\beta_y|\overline{X},C) \end{aligned}. \quad \textbf{[S20]}$$

Because the nonsocial asymmetry can modulate the probabilities for the behaviors, these relations need not be satisfied exactly. However, this effect is probably much weaker than the effect of the nonsocial asymmetry on the nonsocial term in Eq. **S19**. Therefore, for simplicity we use relations (Eq. **S20**) also for the asymmetric setup. The good fit with experimental data confirms that they are a good approximation.

According to Eqs. **S20** and **S9**, we have that $s_{xx} = s_{yy}$ and $s_{xy} = s_{yx}$, and using the definitions in Eq. **S16**, we find that Eqs. **S11** and **S12** become

$$\begin{aligned} P(X|C,B) &= \frac{1}{1 + a_x s^{-(n_x - k n_y)}} \\ P(Y|C,B) &= \frac{1}{1 + a_y s^{-(n_y - k n_x)}} \end{aligned}. \quad \textbf{[S21]}$$

The probability of choosing x or y is then obtained using probability matching (Eq. 2) to get

$$P_x = \left(1 + \frac{1 + a_x s^{-(n_x - k\, n_y)}}{1 + a_y s^{-(n_y - k\, n_x)}}\right)^{-1}, \quad \textbf{[S22]}$$

represented in Fig. S13.

Derivation of an Equation for a Symmetric Setup but with Different Types of Animals to Follow. When there are different types of animals to follow, as in Fig. 4*A Bottom* row, following the steps of *Derivation of Eq. 1*, we find that each type of animal has its own reliability s. For the particular case of the experiment of ref. 2, we have three different types of animals (real animals, the most attractive replica, and the less-attractive replica, with reliability parameters s, s_R, and s_r, respectively). When the most attractive replica goes to x and the less attractive one goes to y, Eq. **S18** becomes

$$P_x = \left(1 + \frac{1 + a s^{-(n_x - k\, n_y)} s_R^{-1} s_r^k}{1 + a s^{-(n_y - k\, n_x)} s_R^k s_r^{-1}}\right)^{-1}. \quad \textbf{[S23]}$$

Demonstration That the Model in Ref. 1 Is a Particular Case of Eq. 3. The decision-making model we used in ref. 1 was developed for a case in which an animal has to choose using the probability that an option is the best one, whereas the model in this paper is for estimated good options. In ref. 1 we obtained that the probability of choosing x in a two-choice setup that can present an asymmetry as

$$P_x = \left(1 + a_{\text{old}}\ s^{-(n_x - n_y)}\right)^{-1}, \quad \textbf{[S24]}$$

with $a_{\text{old}} = 1$ for the symmetric case.

Multiplying and dividing inside the brackets of Eq. **S24** by $\left(1 + \frac{1}{a_{\text{old}}} s^{-(n_y - n_x)}\right)$, we rewrite this expression as

$$P_x = \left(1 + \frac{1 + a_{\text{old}}\ s^{-(n_x - n_y)}}{1 + a_{\text{old}}^{-1} s^{-(n_y - n_x)}}\right)^{-1}, \quad \textbf{[S25]}$$

so that Eq. **S22** reduces to Eq. **S24** for

$$\begin{aligned} k &= 1 \\ a_x &= a_y^{-1} = a_{\text{old}}\,, \end{aligned} \qquad \textbf{[S26]}$$

as we wanted to demonstrate.

Derivation of an Expression for the Point τ Separating the Low-Number and High-Number Decision Behaviors, and Proof of the Approximate ΔN Rule for Low N. We now consider the general expression of the probability (Eq. **S22**)

$$P_x = \left(1 + \frac{1 + a_x s^{-(n_x - k\,n_y)}}{1 + a_y s^{-(n_y - k\,n_x)}}\right)^{-1}. \qquad \textbf{[S27]}$$

For the reasons described below, the transition between the two regimes takes place when the following conditions are met

$$a_x s^{-(n_x - k\,n_y)} = 1 \text{ and } a_y s^{-(n_y - k\,n_x)} = 1. \qquad \textbf{[S28]}$$

These conditions define a point (τ_x,τ_y) with (Fig. S13):

$$\tau_x = \frac{\log(a_x) + k\log(a_y)}{(1-k^2)\log(s)} \qquad \textbf{[S29]}$$

$$\tau_y = \frac{k\log(a_x) + \log(a_y)}{(1-k^2)\log(s)}.$$

This transition point is relevant because when the left-side terms of Eq. **S28** are much lower than 1 they can be neglected, so P_x is always 0.5. Therefore, the region above the transition point (τ_x,τ_y) in which both left-side terms of Eq. **S28** are lower than 1 (region 1 in Fig. S14) is the plateau of $P_x = 0.5$.

However, if the two left-side terms of Eq. **S28** are much higher than 1, we can use the approximations $1 + a_x s^{-(n_x - k\,n_y)} \approx a_x s^{-(n_x - k\,n_y)}$ and $1 + a_y s^{-(n_y - k\,n_x)} \approx a_y s^{-(n_y - k\,n_x)}$ to write Eq. **S27** as

$$P_x \approx \left(1 + a_x/a_y s^{-\Delta N(1+k)}\right)^{-1}, \qquad \textbf{[S30]}$$

which only depends on ΔN. Therefore, the region below the transition point (τ_x,τ_y) in which both left-side terms of Eq. **S28** are higher than 1 (region 2 in Fig. S14) corresponds to a ΔN rule for decision making.

For the case of symmetric nonsocial information, in which $a \equiv a_x = a_y$, Eq. **S29** reduces to

$$\tau \equiv \tau_x = \tau_y = \frac{\log(a)}{(1-k)\log(s)}. \qquad \textbf{[S31]}$$

1. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.

2. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.

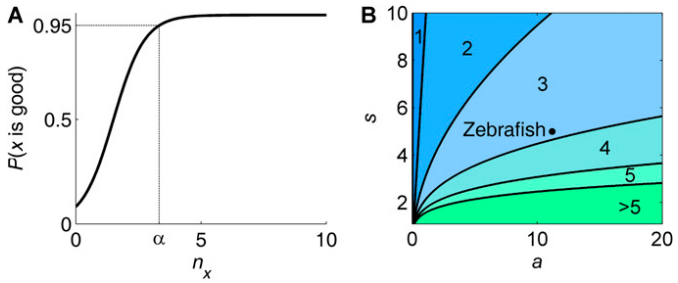


Fig. S1. Maximum number of individuals (α) that need to be counted according to the model for $k = 0$. (*A*) Probability that option x is good (Eq. **1**) here plotted for parameters $a = 11.2$, $s = 5$, and $k = 0$. For $k = 0$, this probability only depends on the variable n_x, increasing as n_x increases to a value of 1. We compute α as the value of n_x for which the probability in Eq. **1** reaches 0.95, getting $\alpha = (\log(a) + \log(1/0.95 - 1))/\log(s)$. Because for $k = 0$ the probability to choose x (Eq. **3**) only depends on n_x through P (x is good), to make the decision the animals do not need to keep count of n_x beyond α. (*B*) The number up to which animals need to count, α, as a function of parameters a and s. For the parameters of the zebrafish dataset, animals only need to count up to 3.

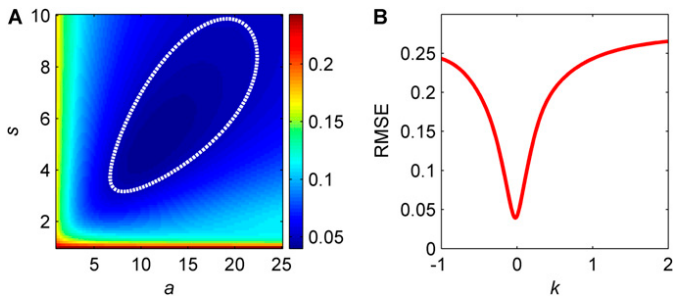


Fig. S2. Robustness of fit to zebrafish data. (*A*) Root mean-squared error between model predictions and data as a function of a and s ($k = 0$). The dotted line limits the region with error below 0.05. (*B*) Root mean-squared error between model and data as a function of k ($a = 11.2$, $s = 5$).

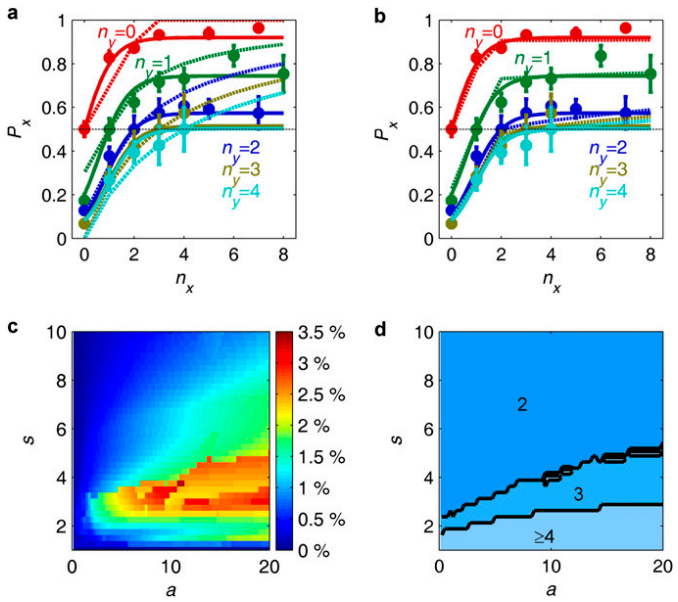


Fig. S3. A simple mechanistic model gives an approximation to Eq. **3**. (*A*) Comparison between decisions using a simple stochastic model (dashed lines) and the model in Eq. **3** (solid lines). In the stochastic model, the focal fish either follows one of the other fish present in the setup (going to the zone where the followed fish is) or does not follow anyone (and therefore moves randomly). If there are N fish in the setup (apart from the focal one), the focal fish will follow any of them with equal probability P when $NP < 1$ and otherwise with probability $1/N$. The probability of not following another fish, and thus choosing at random, is then $\max(\{1 - NP, 0\})$. We modeled the experiment as a series of repeated decisions following this rule, and calculated the time spent at each side in the limit of infinite decisions. Despite the simplicity of this simple stochastic model, it already shows some of the qualitative features of the data. (*B*) Same as *A*, but now the stochastic model considers that the focal fish has a different probability to follow close and far individuals. The implementation of the model was as follows. The probability of not following anyone is now $\max(\{(1 - N_{close}P_{close} - N_{far}P_{far}), P_{nF}\})$, where N_{close} (N_{far}) is the number of fish in the same (opposite) zone as the focal fish, and P_{nF} is the minimum probability of not following anyone. When $N_{close}P_{close} + N_{far}P_{far} > 1 - P_{nF}$, P_{close} and P_{far} are renormalized so that $N_{close}P_{close} + N_{far}P_{far} = 1 - P_{nF}$, whereas P_{close}/P_{far} remains constant. The model with $P_{close} = 0.71$, $P_{far} = 0.005$, $P_{nF} = 0.1$ (dashed lines) has a very good agreement both with the model in Eq. **3** (solid lines) and the experimental data (points). (*C*) Difference between the model in Eq. **3** and the mechanistic model in *B* as a function of a and s. For most values, there is a close agreement. (*D*) Maximum number of individuals that is necessary to count according to model in *B* when parameters are fitted to match the model in Eq. **3**. For most parameter values, we can make $P_{far} = 0$ without a significant worsening of the fit. Then, the probability of not following any fish is $\max\{(1 - N_{close}P_{close}), P_{nF}\}$, which saturates when $N_{close} \geq (1 - P_{nF})/P_{close}$. Due to this saturation, the fish only needs to count up to $(1 - P_{nF})/P_{close}$. This model is consistent with the notion that for a very wide parameter range, animals only need to count up to a small number.

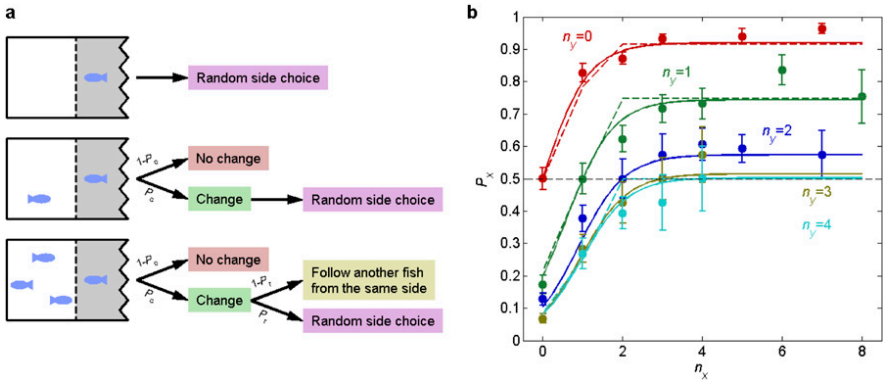


Fig. S4. A very simple mechanistic model gives an approximation to Eq. **3** for parameters corresponding to zebrafish data. (*A*) Schematic diagram of the model. The focal fish (the one on the shaded area) takes into account only the fish that are at the same side. If there are no other fish at the same side, the focal fish moves randomly, and therefore has probability one-half of choosing any side at the next time step (*Top*). If there are other fish at its side, the focal fish follows one of them; at the next decision, it chooses either to stay following the same fish (with probability $1 - P_c$) or to change (with probability P_c) and follow another fish, or not follow anyone. If there is only one fish at the same side, changing means necessarily not following anyone in the next time step, and therefore moving randomly (*Middle*). If there are more than one fish, then changing may lead to follow another fish and therefore remain at the same side, with probability $1 - P_r$, or not follow anyone, with probability P_r (*Bottom*). (*B*) Comparison among model in *A* (dashed lines), model in Eq. **3** (solid lines), and experimental data (points) for $P_c = 0.28$ and $P_r = 0.34$. The correspondence is good except for the $n_y = 2$ case (blue). The model corresponds to "counting up to 2", whereas the data are best fitted with a "counting up to 3" model, as in the more complex model of Fig. S3*B*.

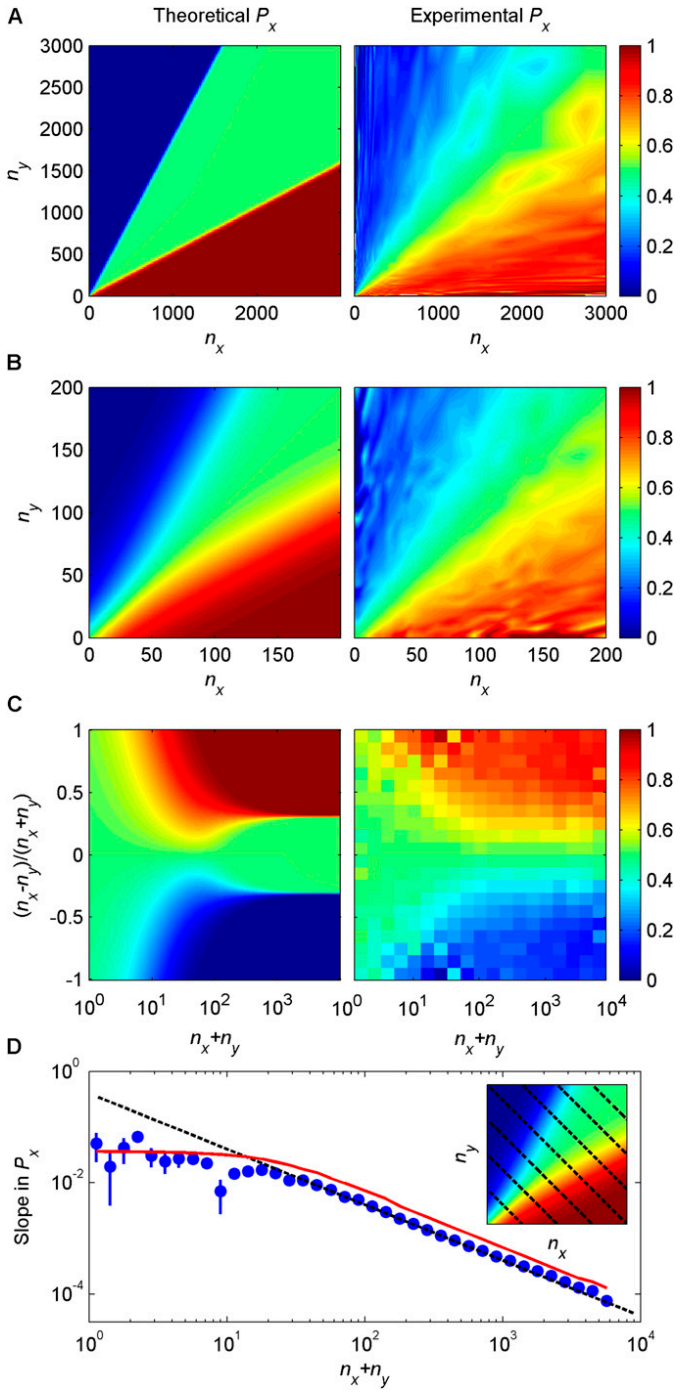


Fig. S5. Eq. **3** reproduces the structure of the ant dataset. Same as Fig. 3 but comparing the ant dataset to Eq. **3** (or, equivalently, Eq. **4** setting $p_{rand} = 0$ instead of the value $p_{rand} = 0.39$ in the main text). (*A*) Probability of choosing option *x* as a function of how many ants have previously been at locations *x* and *y*, n_x and n_y, for theory (*Left*) using Eq. **4** with $a = 2.5$, $s = 1.07$, $k = 0.53$, $p_{rand} = 0$ and experiments (*Right*). (*B*) Detail of *A*. (*C*) Same as *A* but represented as a function of $\Delta N/N$ and *N*. (*D*) Slope of the probability of choosing *x* in *A* as obtained from a linear fit along the lines depicted in the *Inset*. Experimental values (blue dots; error bars are 95% confidence interval), theory (red line), and Weber law (black line).

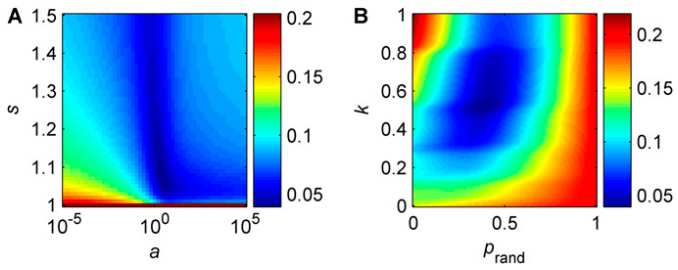


Fig. S6. Robustness of the fit of the model in Eq. **4** to the ant dataset. (*A*) Mean squared error between model and data as a function of parameters a and s, for $k = 0.53$ and $p_{rand} = 0.39$. To adequately sample the data, that span several orders of magnitude, we scanned the $n_x - n_y$ plane using sections of constant $n_x + n_y$ equispaced in a logarithmic scale, instead of a square grid. (*B*) Mean squared error as a function of k and p_{rand} for $a = 2.5$ and $s = 1.07$. Sampling of the $n_x - n_y$ plane as for *A*.

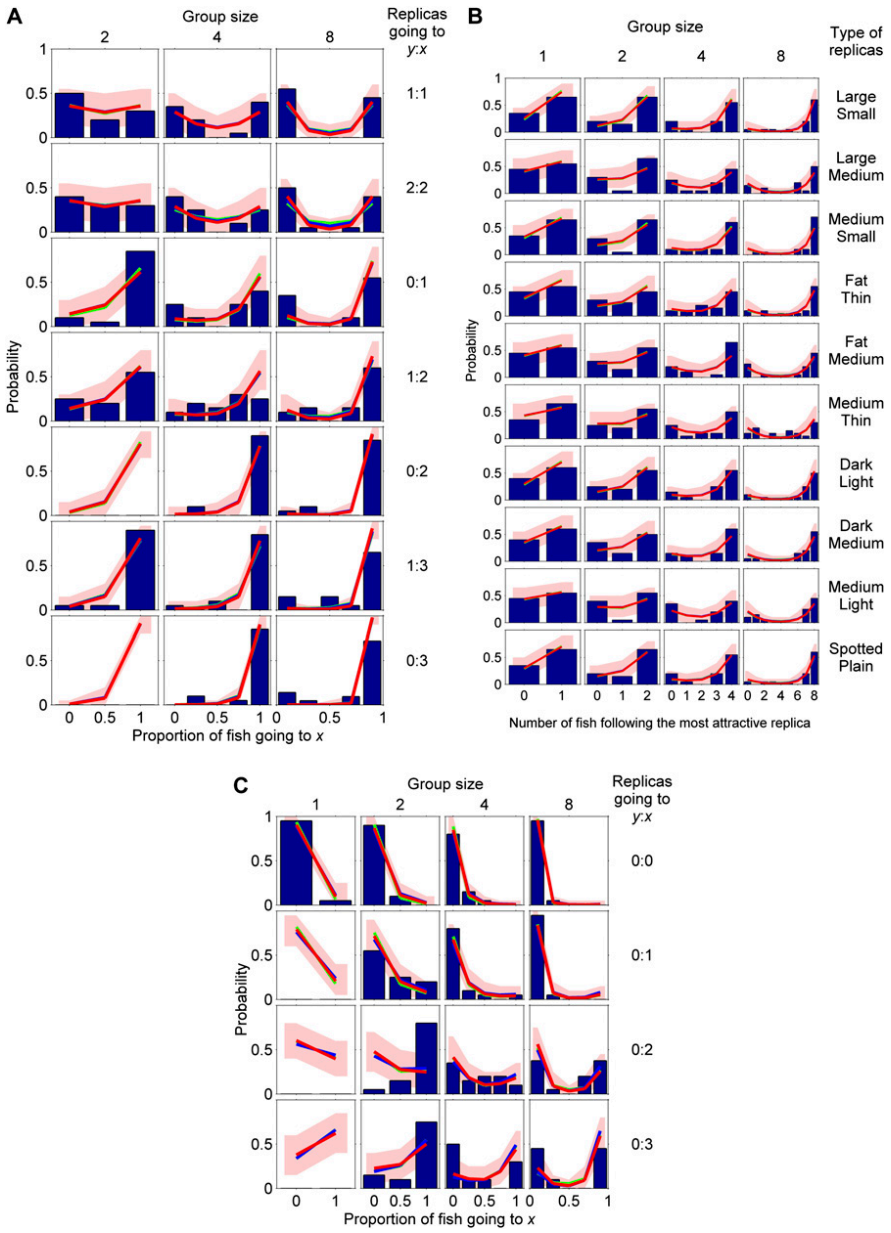


Fig. S7. Complete stickleback dataset (1, 2) and model fits. The three figures show experimental data as blue histograms and results for the $k = 1$ model (3) as red lines and green and blue lines for $k = 0.5$ and $k = 0$, respectively. In the three cases, $s = 2.5$ and a was refitted for each k. Pink regions limit the 95% confidence intervals for the $k = 1$ case. (*A*) Results for symmetric setup with different number of replica fish going to each side (e.g., 1:2 means one replica going to y and two replicas going to x). $a_x = a_y = 1$ for $k = 1$ (red line), $a_x = a_y = 5$ for $k = 0.5$ (green line), and $a_x = a_y = 224$ for $k = 0$ (blue line). (*B*) Results for symmetric setup and differently modified replica fish going to each side. We set the intermediate replica's reliability parameter equal to the one of the real fish ($s = 2.5$), and adjust the others to match the ratios found in ref. 3. We got $s_{small} = 1.25$, $s_{medium} = 2.5$, $s_{large} = 3.57$, $s_{thin} = 1.88$, $s_{medium} = 2.5$, $s_{fat} = 3.62$, $s_{light} = 1.95$, $s_{medium} = 2.5$, $s_{dark} = 4.55$, $s_{plain} = 2.5$, and $s_{spotted} = 5.81$. Parameter a as in *A*. (*C*) Results for setup with a replica predator at x. $a_x = 9.5$, $a_y = 1/9.5$ for $k = 1$ (red line), $a_x = 1.25$, $a_y = 31.5$ for $k = 0.5$ (green line), $a_x = 1{,}250$, and $a_y = 10{,}000$ (in this case, if we multiply these two parameters by any number greater than 0.1, the fit changes very little) for $k = 0$ (blue line).

1. Ward AJW, Sumpter DJT, Couzin ID, Hart PJB, Krause J (2008) Quorum decision-making facilitates information transfer in fish shoals. *Proc Natl Acad Sci USA* 105(19):6948–6953.
2. Sumpter DJT, Krause J, James R, Couzin ID, Ward AJW (2008) Consensus decision making by fish. *Curr Biol* 18(22):1773–1777.
3. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.

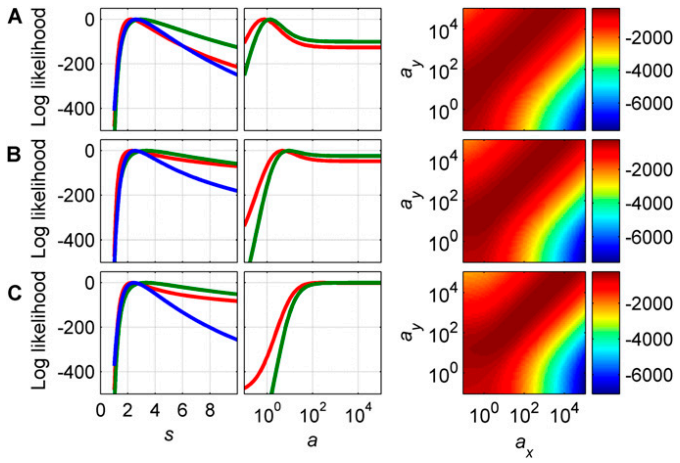


Fig. S8. Robustness of the fit for the stickleback dataset. (*A*) Results of fits for $k = 1$. (*Left*) Log-likelihood as a function of parameter s (symmetric setup in red, setup with two different types of replicas in green, setup with predator in blue). (*Center*) Log-likelihood as a function of parameter a (red for symmetric setup and green for setup with modified replicas). (*Right*) Log-likelihood as a function of parameters a_x and a_y for the asymmetric setup with predator. (*B*) Same as *A* but for $k = 0.5$. (*C*) Same as *A* but for $k = 0$. All log-likelihoods are relative to their maximum value.

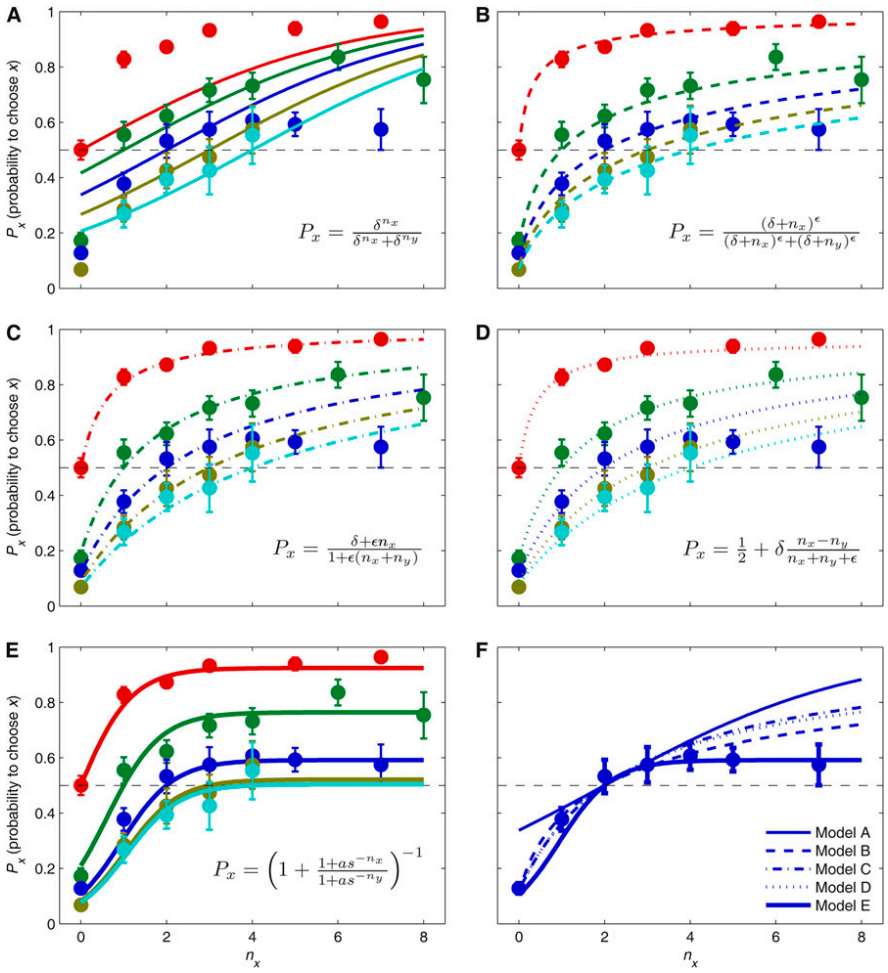


Fig. S9. Best fit of different functions to zebrafish dataset. (*A*) Logistic function $P_x = \delta^{n_x}/(\delta^{n_x}+\delta^{n_y})$, as in refs. 1 and 2, for $\delta = 1.4$. (*B*) $P_x = (\delta+n_x)^\varepsilon/((\delta+n_x)^\varepsilon+(\delta+n_y)^\varepsilon)$, as in refs. 3 and 4, for $\delta = 0.1$ and $\varepsilon = 0.7$. (*C*) $P_x = (\delta+\varepsilon n_x)/(1+\varepsilon(n_x+n_y))$, as in ref. 5, for $\delta = 0.5$ and $\varepsilon = 1.6$. (*D*) $P_x = 0.5+\delta(n_x-n_y)/(n_x+n_y+\varepsilon)$, as in ref. 6, for $\delta = 0.48$ and $\varepsilon = 0.47$. (*E*) Our model in Eq. **3** for $a = 11.2$, $s = 5$. (*F*) Comparison of the five previous models for line $n_y = 2$. Only the model in Eq. **3** gives a good fit in this region.

1. Ward AJW, Herbert-Read JE, Sumpter DJT, Krause J (2011) Fast and accurate decisions through collective vigilance in fish shoals. *Proc Natl Acad Sci USA* 108(6):2312–2315.
2. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
3. Goss S, Aron S, Deneubourg JL, Pastels JM (1989) Self-organized shortcuts in the Argentine ant. *Naturwissenschaften* 76:579–581.
4. Deneubourg J, Aron S, Goss S, Pasteels J (1990) The self-organizing exploratory pattern of the argentine ant. *J Insect Behav* 3(2):159–168.
5. Meunier H, Leca JB, Deneubourg JL, Petit O (2006) Group movement decisions in capuchin monkeys: The utility of an experimental study and a mathematical model to explore the relationship between individual and collective behaviours. *Behaviour* 143(12):1511–1527.
6. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.

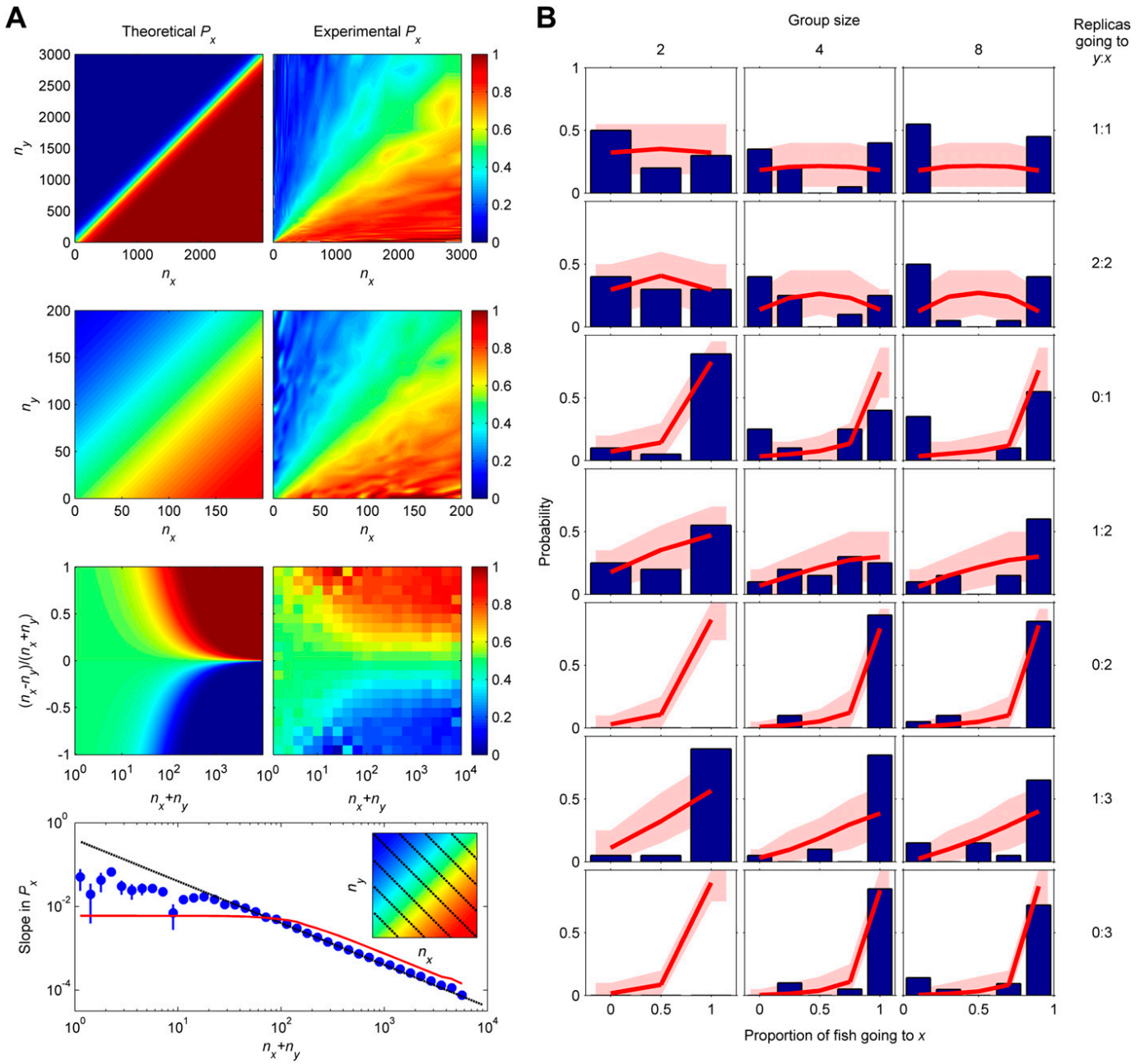


Fig. S10. The models in refs. 1 and 2 do not explain other datasets. (*A*) Same as Fig. 3, but using $P_x = (1 + s^{-(n_x - n_y)})^{-1}$, with $s = 1.012$. This model was used in ref. 1 to describe the stickleback dataset, and cannot describe the ant dataset. (*B*) Same as Fig. S7*A*, but with $P_x = 0.5 + A(n_x - n_y)/(n_x + n_y + T)$, with $A = 0.5$ and $T = 0.4$, which is the function used in ref. 2 to describe the ant dataset, with the 0.5 term added and with A restricted between 0 and 0.5, so that probabilities are between 0 and 1.

1. Pérez-Escudero A, de Polavieja GG (2011) Collective animal behavior from Bayesian estimation and probability matching. *PLOS Comput Biol* 7(11):e1002282.
2. Perna A, et al. (2012) Individual rules for trail pattern formation in Argentine ants (*Linepithema humile*). *PLOS Comput Biol* 8(7):e1002592.

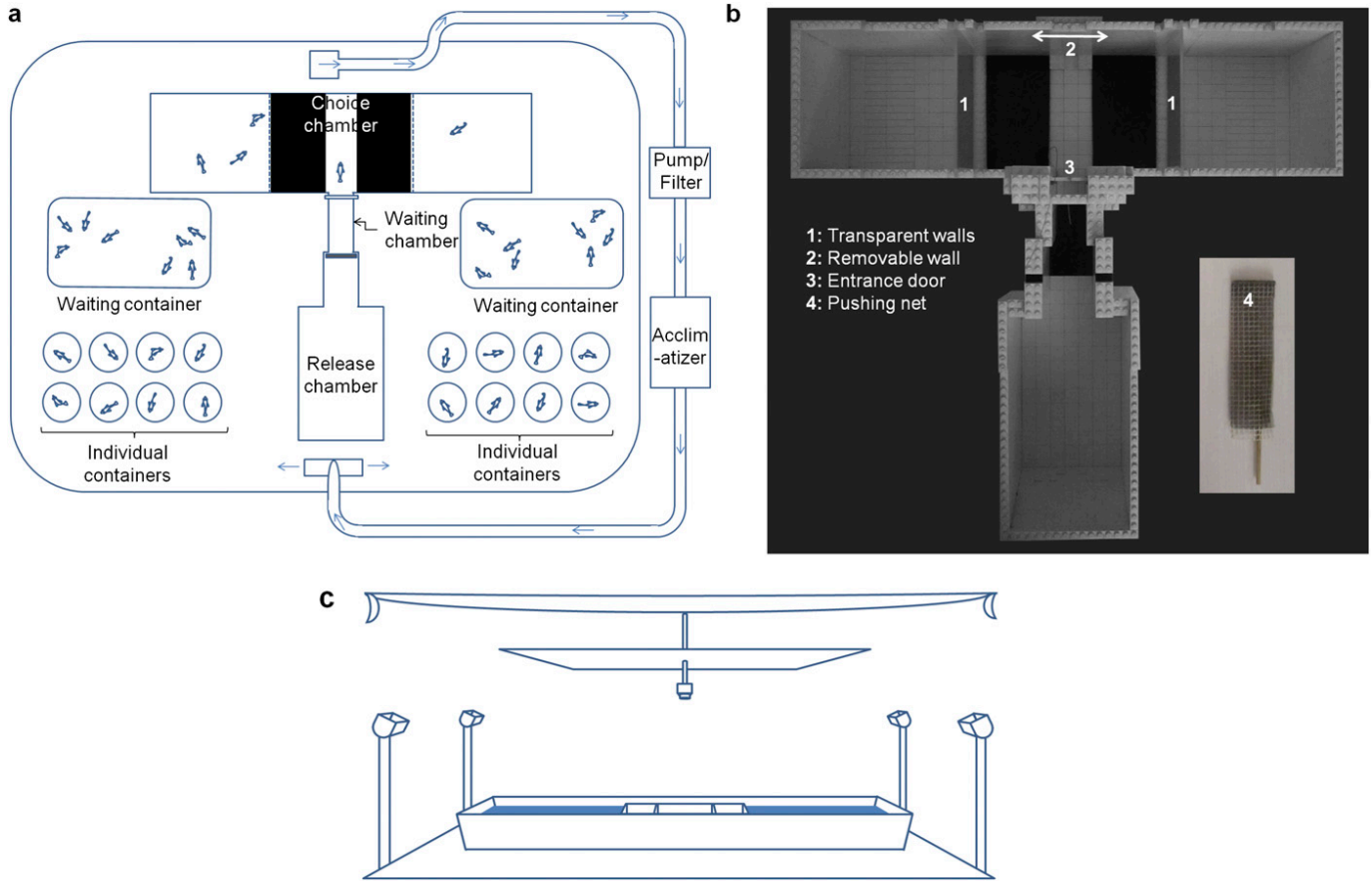


Fig. S11. Experimental setup for zebrafish. (*A*) The behavioral setup is inside a bigger tank so that fish are acclimatized to the same water for 1 d before the experiment, housed in waiting containers in groups of 8–10 fish. At 1 h before the experiment, each fish is isolated and fed with frozen artemia in an individual container. The fish stays in the individual container until placed in the release chamber and gently pushed into the waiting chamber with a net that fits tightly between the walls to prevent the fish from going back to the release chamber. The door to the setup is then lifted and, once the fish enters the setup, it is closed. The camera records for 5 min from the opening of the door. After the experiment, the fish is pushed back to the release chamber, where it is caught. Then, a segment of wall opposite to the entrance door is removed, and water from outside is pumped into the central chamber so that odors are washed out. (*B*) The T-shaped setup is made of white LEGO bricks, with transparent walls separating the three chambers made of UV-transparent PLEXIGLAS (PLEXIGLAS GS 2458; Evonik Para-Chemie). The setup's central chamber (choice chamber) measures 20 × 13 cm. The floor of this central chamber has a central white zone 5 cm wide, and two black lateral zones 7.5 cm wide each. The two lateral chambers measure 14 × 13 cm each. Walls are 17-cm high, and water level is 6 cm. (*C*) Illumination is provided by four 500-W halogen lamps pointing to a white sheet on the ceiling. A Basler A622f camera records from above. An opaque roof just above the camera provides uniform shading on the setup.

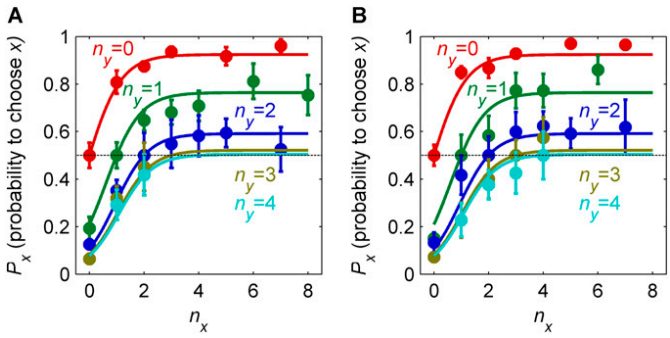


Fig. S12. Comparison of results using naive and nonnaive zebrafish. (*A*) Results for naive zebrafish, which have never seen the setup before the experiment. (*B*) Results for zebrafish that have been tested several times in the setup. Lines correspond to the theoretical model (Eq. **3**) with same parameter values as for Fig. 2*C*.

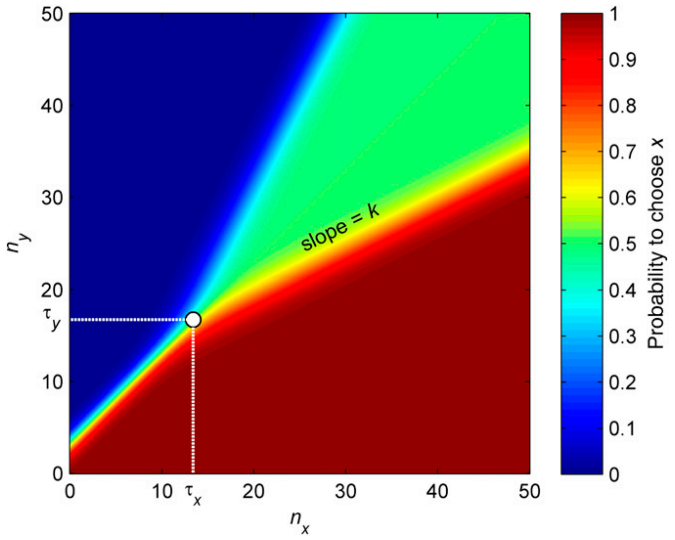


Fig. S13. Probability to choose x, P_x, for the general case of asymmetric nonsocial information (Eq. **S22**). Parameters: $s = 2.5$, $k = 0.5$, $a_x = 100$, $a_y = 10{,}000$. Compare this figure with the one corresponding to symmetric social information in Fig. 1*B*. See Eq. **S29** for an analytical expression of (τ_x, τ_y).

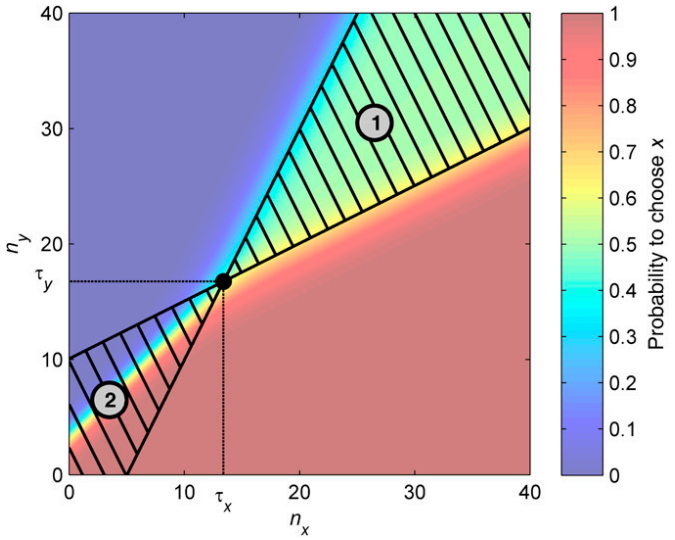


Fig. S14. Transition point τ between the low- and high-numbers regimes. Region 1 corresponds to the plateau with $P_x = 0.5$. The ΔN rule is approximately valid in region 2. Parameters are as in Fig. S13.

6
Tracking & Identification

6.1 *Introduction*

All animals are equal, but some animals are more equal than others.

George Orwell
Animal Farm

An important datasets in the study of the behavior of a group of animals is the set of their trajectories. Although it may seem trivial, obtaining these trajectories automatically is in practice quite difficult, and we have not found any state-of-the-art system capable of tracking small laboratory animals reliably with sufficient temporal and spatial resolution.

Our aim was to be able to extract the trajectories of the animals of a group while they are moving in a controlled environment in the laboratory (for example, fish in a tank, or mice in a cage). Therefore, we have the advantage of being able to control the conditions of the set-up (illumination, etc.). Also, we restricted ourselves to the case in which animals move essentially in 2 dimensions, either because it is their natural way of moving (e.g. mice) or because we restrict them to do so (e.g. fish in shallow water). While these controlled conditions facilitate our task, we had very stringent requirements: The program should be able to keep the correct identity of each individual during long videos, and should be 100% automatic (i.e. not require manual correction).

Current techniques based on attaching a tracking device to each individual, such as GPS loggers (Nagy et al., 2010; King et al., 2012) or RFID transponders (Streit et al., 2003; Robinson et al., 2008; Molet et al., 2008; Lewejohann et al., 2009) are not adequate for our conditions: GPS loggers have an accuracy of the order of one meter and a temporal resolution of the order of the second, and do not work well inside buildings. They are therefore useful for big animals that move long distances and whose trajectories do not change too fast, not for laboratory animals. RFID transponders can be much smaller, being usable even with insects (Streit et al., 2003; Robinson et al., 2008; Molet et al., 2008). But current technology does not allow the tracking of RFID transponders from far away. They are therefore ideal to detect the passage of an animal through a certain region, but current tracking devices based on RFID have too low temporal and/or spatial resolution.

Besides Gonzalo G. de Polavieja (director of this thesis) several other members of the laboratory have contributed to the work presented in this chapter, especially Sara Arganda Carreras, Julián Vicente Page and Robert Hinz. Some of the experimental set-ups described here have been built by them, and some results and images presented in this chapter come from videos recorded by them.

An alternative to these techniques is to record the animals in a video, and then extract the trajectories from it. Video has the advan-

tage of having high spatial and temporal resolution, not requiring manipulation of the animals to attach any tracking devices, and being relatively inexpensive. However, it is not easy to extract from the video the trajectory of each individual. In laboratory conditions, with uniform illumination and uniform background the images of the animals are easy to separate from the background. In these favorable conditions it is easy to track one isolated individual, just extracting its image from every frame and making a sequence with all of them. But when several animals move together, it is difficult to follow each individual separately, without mixing its identity with those of the other individuals. In particular, when two or more individuals cross and their images overlap during several frames as in Figure 6.1, it is difficult to assign each individual after the crossing to the identities before the crossing. Note that we are mainly interested in highly social species, for which it is very frequent to observe many individuals overlapping in a small area (for example as in Figure 6.2), so we need to deal with crossings of very high complexity.

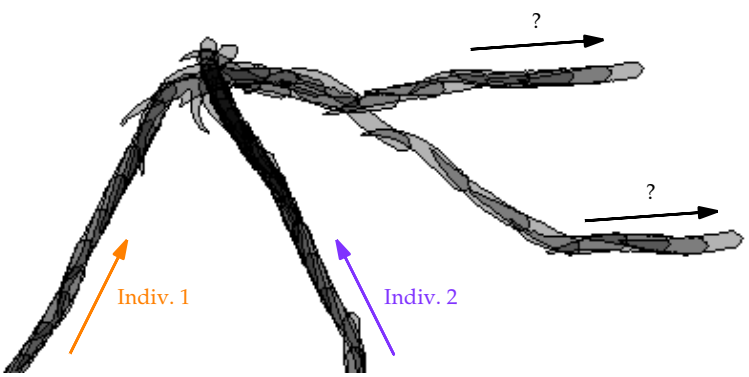


Figure 6.1: Silhouettes of two zebrafish crossing. The two fish have identities 1 and 2 before the crossing. While the individuals are separated, it is easy to keep track of each individual's identity. But when the two animals overlap the tracking program loses the identities, being difficult to re-assign them correctly afterwards.

Several different techniques have been tried to resolve the crossings (see reviews for example in Gabriel et al., 2003; Delcourt et al., 2012).[1]

One source of information is the movement of each individual before and after the crossing. Thus, one can find the assignment of identities that links the dynamical state of each individual before and after the crossing in a way most compatible with the typical movement of animals. A basic version of this approach consists of taking the center of mass of each individual in the last frame before the crossing and its velocity at that point, and extrapolate assuming that it kept constant velocity during the crossing. After the crossing, one relates each individual to the one whose extrapolated position is nearest to it (Branson et al., 2009; Delcourt et al., 2009). Instead of using the center of mass, that is sensitive to segmentation errors, one can use the information of the full image, computing its optical flow (Sanchez and Dibos, 2004). One can also include information about the velocity not only before the crossing but also after it, as well as accelerations, orientation and other dynamical parameters. It is also useful to take into account the statistics of how the animals

[1] In the following paragraphs we will discuss each technique separately, but they can also be used in combination. There are powerful mathematical techniques to integrate the information provided by each of them, such as Kalman filters or particle filters (Okuma et al., 2004; Sanjeev Arulampalam et al., 2002; Khan et al., 2005)

move in other parts of the video. These techniques may obtain very good results in the case of objects whose movement is very smooth (for example, cars or planes). This is not the case of most animals, that very often change direction suddenly.[2]

Besides the dynamical information extracted from the path before and after the crossing, additional information can be extracted from the images in which the individuals overlap. Some methods repeat the segmentation of these images, trying to obtain a separate image of each of the overlapping individuals. This may be achieved for example by using a more restrictive threshold (Branson et al., 2009), eroding the original image until two or more blobs emerge (Kato et al., 2004) or dividing the blob into pieces whose size and shape match those of the isolated individuals (Branson et al., 2009). This approach may be quite effective when the individuals are not overalpping completely,[3] but will fail very often when the occlusion is significant.[4] The system can use the images of the overlapping objects to extract additional information, such as who is in front and who is at the back (McKenna et al., 2000), but to achieve high reliability these methods normally need more difference in the textures of the two objects than the one existing between unmarked individuals of the same species.

A further sophistication consists of building a model of the shape of each individual (Branson and Belongie, 2005; Fontaine et al., 2008; de Chaumont et al., 2012; Butail and Paley, 2012). The images of overlapping animals can be then interpreted by fitting several overlapping shape models to them, so that each animal can be tracked through the crossing. This technique has the disadvantage that the model is valid only for one species (or several species with very similar shape and that move similarly). In exchange, in theory if the model is well constructed and the images have enough spatial and temporal resolution, this technique may be able to unambiguously resolve a very high proportion of crossings. In practice, however, it is very difficult to construct a model that takes into account all possible deformations of the animal with enough accuracy, and when several animals overlap it is very difficult to obtain a robust fitting of the models to the image. Therefore, these systems are very useful to better characterize the behavior of the animals when they are separated, but they cannot resolve complex crossings[5]

Another alternative is the use of several cameras pointing from different angles in order to reconstruct the 3D trajectories of the animals (Butail and Paley, 2012; Attanasi et al., 2013). With several cameras it is less likely to have some animals overlapping in the images of all cameras at the same time, and this redundancy can be used to resolve many crossings (Attanasi et al., 2013). But, especially for high numbers of animals, some simultaneous overlappings do occur, and these systems also lose some individual identities eventually.[6] Also, the extra complexity of the experimental set-up limits the applicability of these techniques; current implementations need very high frame rates (Butail and Paley, 2012; Attanasi et al., 2013), so the

[2] For example, Delcourt et al. (2009) report that this technique resolves correctly between 50% and 83% of crossings in videos of fish (*O. niloticus*).

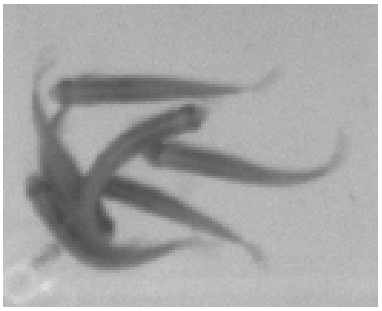


Figure 6.2: Seven zebrafish overlapping in the same area.

[3] Branson et al. (2009) report that a combination of re-segmentation with more stringent threshold and division of blobs by size works very well in walking flies (*D. melanogaster*), that normally overlap very little. A small amount of mistakes persist, that must be corrected manually.

[4] Kato et al. (2004) show that for zebrafish re-segmentation through erosion of the image is not useful for videos with more than 3 individuals, and that even in videos with 2 individuals an important amount of mistakes persist.

[5] Butail and Paley (2012) describe a system that solves crossings involving less than 4 individuals. de Chaumont et al. (2012) report around 2 mistakes per minute in videos of 2 mice. Branson and Belongie (2005) report 64% of correct crossings in a video with three mice.

[6] Attanasi et al. (2013) follow successfully around 90% of the individuals during more of the 90% of the duration of each trial, that may last up to 12 seconds. See also note 5.

video cannot be directly stored in the computer's hard disk, limiting the length of the videos.[7]

Therefore, while current tracking techniques can efficiently extract non-labeled trajectories (Kato et al., 2004; Branson et al., 2009; Delcourt et al., 2009; de Chaumont et al., 2012; Butail and Paley, 2012; Attanasi et al., 2013), can describe the posture of individual animals with a great detail (Fontaine et al., 2008; de Chaumont et al., 2012; Butail and Paley, 2012) and have achieved a very high degree of accuracy in the automatic classification of behaviors (Branson et al., 2009; de Chaumont et al., 2012), no current tracking algorithm is capable of correctly resolving all crossings in a group of animals. For example, in the case of fish we have found no report of success rate higher than 83%. Note that 17% mistakes in the crossings translates to much more than 17% of the trajectories with mistaken identities because once the identities of two individuals have been swapped they remained swapped for the rest of the video, so one single mistake results in a very long portion of the video with mistaken identities. Because in a group of social animals we will typically observe several hundreds of crossings in a few minutes, even with low error rates after some time most individuals will have been affected by at least one mistake, their identities being essentially equivalent to random assignment. For example, Figure 6.3a shows the results that an algorithm capable of resolving correctly 83% of the crossings would obtain on a video of 8 zebrafish (colored lines indicate that the identity is correctly recovered by the algorithm, and black lines indicate mistaken identities). Even if algorithms to resolve crossings improve significantly, it is unlikely that they will be able to deliver reliable results, because of propagation of errors. For example, Figure 6.3b illustrates the results that would be provided by an almost ideal algorithm that solved 99.5% of the crossings. Even with this unrealistically high accuracy, there are large portions of the final trajectories with mistaken identities.

[7] Butail and Paley (2012) and Attanasi et al. (2013) record less than 15 seconds per video.

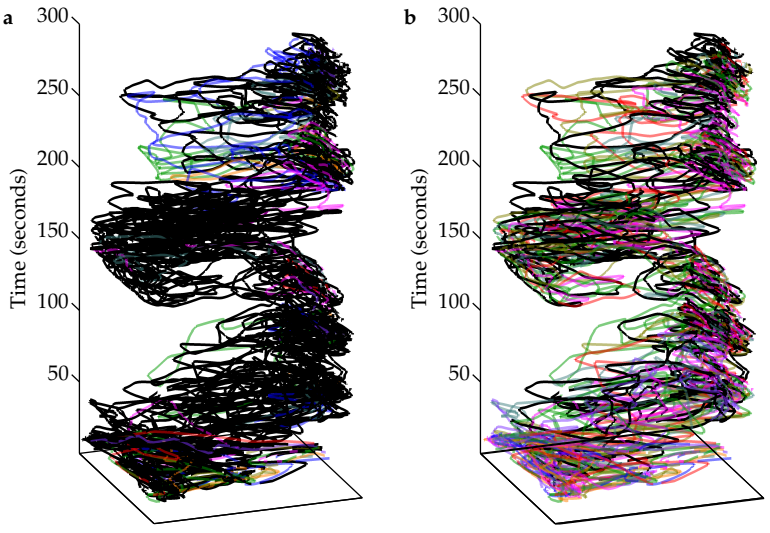


Figure 6.3: Illustration of mistakes made by crossing-solving algorithms. The figures show the trajectories of 8 zebrafish. x and y axes correspond to the x and y coordinates of the fish as they move inside a square set-up (0.5×0.5 m), and the z axis corresponds to time. In order to simulate the outcome of different tracking algorithms, we started from the correct trajectories (validated manually). Then, for each crossing we swapped identities of two of the crossing fish with a given probability, simulating errors in the tracking. In this figure, colored lines correspond to portions of trajectories with correct identification (each color corresponding to one individual, and assuming that identities are correct at the beginning of the video). Black lines correspond to portions with wrong identity. **a.** Simulation of a tracking algorithm that resolves 83% of the crossings correctly. **b.** Simulation of an algorithm that resolves 99.5% of the crossings correctly.

Therefore, current methods for animal tracking are not capable to reliably obtain trajectories with correct identities. One alternative is simply to concetrate on aspects of behavior for which identities are not important, but then we would have to neglect very important features such as individual differences, leadership, aggression, etc. Another alternative is to manually correct the videos. Given the astonnishing performance of human vision, manual correction may actually achieve almost the 100% accuracy needed for the task, and so far all studies that report individual trajectories have used manual correction (Branson et al., 2009; Herbert-Read et al., 2011; de Chaumont et al., 2012). But even humans can make some mistakes, so the risk of having a substantial part of the video with interchanged identities remains. Furthermore, the task of manually correcting the videos is huge, because even if the algorithm that made the original tracking has a high accuracy, say 90%, this does not mean that the human operator will have to review and correct only 10% of the crossings. The operator will need to review every single crossing to make sure that they are correctly resolved, and then correct the 10% that were actually mistaken. Even with specialized software that helps to navigate the video from one crossing to the next and easily correct the mistakes, it is impracticable to do it in a regular basis.[8]

[8] For example, manual correction of the zebrafish trajectories shown in Figure 6.3 required about 50 minutes of correction *per minute* of video. (Branson et al., 2009) reports requiring 6, 24 and 96 minutes of correction per minute of video for videos of 10, 20 and 50 flies, respectively.

Tracking through identification

An ideal way of keeping identities during the whole video is to use physical differences between the individuals to identify them in each frame (Figure 6.4). Using this procedure one can follow each individual along the video, regardless of the complexity of the crossings. Also, in contrast with the methods presented in the previous section in this case errors do not propagate, because making a mistake in a given frame does not affect the identifications of other frames in the video.

Figure 6.4: Silhouettes of two zebrafish that cross, with identifications. The color of each silhouette corresponds to the individual that has been identified in each frame. Black means that the blob was not identified, either because it does not correspond to a single individual (as in the region of crossing), or because no unambiguous identification was possible (black frame in right-hand-side purple trace). Errors do not propagate, because a mistake or lack of identification at one frame does not affect other frames.

The easiest way to identify the individuals is to mark them, for example with tags of different colors (Jerry et al., 2001; Dahlbom et al., 2011; Delcourt et al., 2011). But the marks themselves may affect the results in different ways: First, it requires manipulation

of each individual prior to the experiment, which for some species may be quite difficult. For example, marking zebrafish requires anesthesising them, a procedure that kills a proportion of the individuals. Also, sometimes the marks themselves may affect the behavior. For example, Dennis et al. (2008) report that marked fowls receive more aggression, have lower body mass and have altered levels of epinephrine and dopamine with respect to unmarked animals of the same group. In zebrafish, Dahlbom et al. (2011) report a tendency of fish with blue tags to become dominant in dyadic interactions.

The ideal method would be one capable of identifying each individual using the natural physical differences among them. This is a challenging task, because individuals of the same species are usually very similar to each other. To make matters more difficult, in order to reduce the variability among the individuals many experiments are performed with size-matched siblings. Therefore, the method would need to distinguish individuals that have been selected precisely for being as similar to each other as possible, and that usually are indistinguishable to the human eye. Furthermore, identification must be done while the animal is moving freely, so it must be able to detect the small differences between the images of two individuals, while discarding the big differences that exist between two images of the same individual in different postures (Figure 6.5).

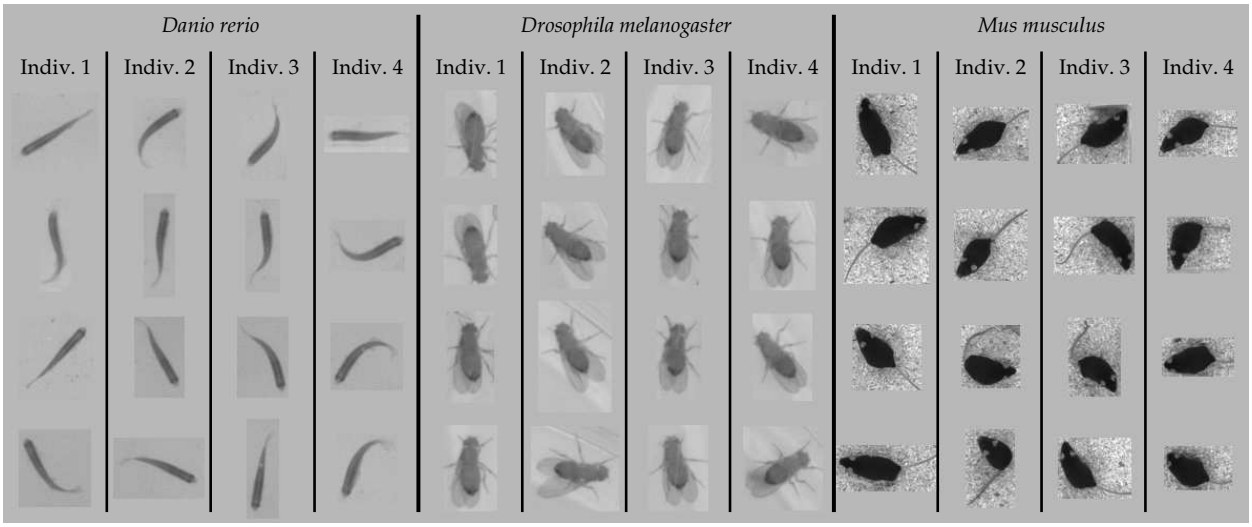


Figure 6.5: Example of variability between images of the same individual and of different individuals, for three different species. Each of the four images in each column corresponds to the same individual in different moments of a video.

Several methods exist to track individuals based on their aspect (Balcells et al., 2005; Takala and Pietikainen, 2007). But these methods have been developed mainly for tracking pedestrians and vehicles, that are relatively dissimilar to each other. Also, emphasis on real-time tracking forces these algorithms to work frame-by-frame, restricting the use they can make of the available information. Likewise, most pattern recognition algorithms usually work with objects that are much more dissimilar to each other (such as letters, or different objects) (Flusser and Suk, 1993; Lowe, 1999). Some techniques are capable of distinguishing similar objects (Kovalev and Petrou, 1996),

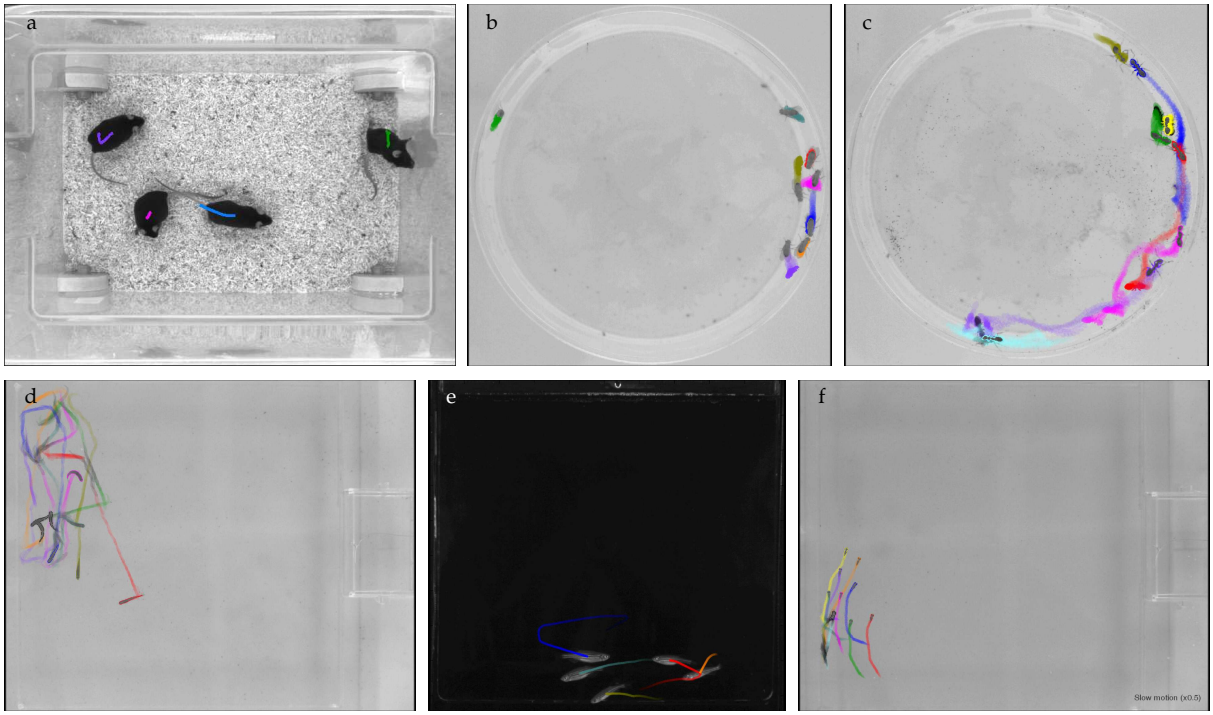


Figure 6.6: The method has been successfully tested in several species. **a.** Two size-matched sibling male mice (*Mus musculus*). **b.** 8 sibling female flies (*Drosophila melanogaster*). **c.** 8 ants (unknown species). **d.** 8 size-matched zebrafish (*Danio rerio*) recorded from above. **e.** 5 zebrafish, recorded from the side. f. 10 medaka fish (*Oryzias latipes*).

but none of them has been applied to tracking, that has the extra requirements of dealing with changes of posture of each individual (Figure 6.5) and of learning the aspect of each individual from the same sequence of video that must be tracked. An exception is human face recognition, for which very powerful techniques exist (Turk and Pentland, 1991; Pentland et al., 1994; Wiskott et al., 1997). But these techniques are not readily applicable to our problem, because some of them use specific features of human faces, and faces do not deform as much as the whole body of animals.

This chapter presents a novel tracking and identification system based on a pattern recognition method capable of distinguishing individuals of the same species, even when they are size-matched siblings.[9] The method is completely general, not relying on any specific feature of a given species. While the individuals of some species may be too similar for the method to distinguish them, we have still not found any such case. We have succesfully tested the method in several different species and conditions, including mice (Figure 6.6a, or see video[10]), flies (Figure 6.6b, video[11]), ants (Figure 6.6c, video[12]), zebrafish recorded from above (Figure 6.6d, video[13]), zebrafish recorded from the side (Figure 6.6e, video[14]), medaka (figure 6.6f, video[15]) and stickleback (video[16]).

Our system is 100% automatic, and does not require fine-tuning of parameters for each video, even when changing animal species and conditions of the set-up[17]. It is not necessary to train the algorithm with images of the individuals before tracking a video. The only input is the file with the video itself, from which the program automatically extracts the information needed to identify each indi-

[9] See youtu.be/MDMRcyJKzdU for a video summary of the system.

[10] youtu.be/yZxepx43h5Q
[11] youtu.be/4zoNCeHksJA
[12] youtu.be/PLZRkneUcS4
[13] youtu.be/l32USFoglC8
[14] youtu.be/UeKqcZUK168
[15] youtu.be/Yk-HNBI21Ns
[16] youtu.be/T3jLL-uxnlk

[17] The only information that must be input by the user is the number of individuals present and two simple parameters for the segmentation: Intensity threshold, and minimum size of the animals. These two parameters have default values that are adequate for 80-90% of the videos we have recorded.

vidual.

The algorithm provides trajectories with correct identities with very high reliability, regardless of the complexity of the crossings.[18] Identification mistakes may occur in some frames but they never propagate, their effect being always limited to very small portions of the video. Figure 6.7 shows the same trajectories as Figure 6.3, but tracked with our system. Only 0.5% of the trajectories are affected by mistakes, in contrast with the case of state-of-the-art algorithms based on solving crossings, for which the most part of the trajectories have mistaken identities (Figure 6.3).

[18] See video in youtu.be/l32USFoglC8 showing the performance of our system for the crossing shown in Figure 6.2

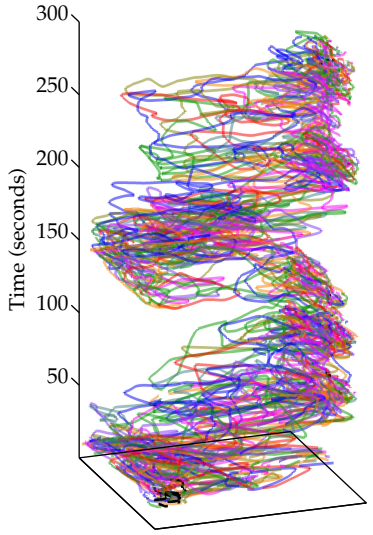


Figure 6.7: Trajectories of 8 zebrafish, as obtained by our tracking algorithm. x and y axes correspond to the x and y coordinates in the experimental set-up, and z axis is time. Each color corresponds to one individual for the portions in which identity is correct. Black trajectories correspond to portions where identity is incorrect (with respect to a manually corrected video) Identities were mistaken only in small intervals, amounting less than 0.5% of the whole trajectories.

6.2 *Description of the method*

This section has been removed because of confidentiality issues.

6.3 *Description of the software*

The software is coded in MATLAB R2010a (The MathWorks, Inc., Natick, Massachusetts, United States), with some small modules compiled in C to improve the performance of the heaviest tasks. We have developed a graphical interface that allows the user to input easily the parameters needed for the tracking (Figure 6.8). The user must enter the number of animals that take part in each video. Also, sometimes it is necessary to tune the segmentation threshold, and the minimum size of blobs. In addition to these three parameters, the software allows the user to input a region of interest (ROI) that may be rectangular or circular. It is also possible to remove rectangular and circular patches from the ROI. Also, the user can choose whether to use background substraction or not, and can choose to invert contrast (in order to switch between conditions in which the animals are darker than the background and conditions in which the animals are lighter than the background). The user can also define a

specific interval of the video to be tracked, instead of the full video. Also, resolution of the video may be decreased (this is useful to accelerate the tracking when the animals are very big with respect to the image). Finally, the maximum number of frames used to build the references may be changed (decreasing this number also accelerates the tracking).

The program can read video in any format supported by the codecs installed in the computer to which MATLAB has access. It can work both with color and monochrome video, but it does not make use of the color information, transforming each color frame to grayscale before processing it.

The software can run in a conventional computer. It requires at least 4 GB RAM, being preferrable to have 8 GB RAM (with 4 GB it is not possible to use parallel computing, and may give problems with more than 8 individuals). The tracking is performed off-line, and typically takes significantly longer than the duration of the video. In a commercial[19] desktop computer with 8 GB RAM and an 8-cores Intel i7 2.67 GHz processor, typically the segmentation and building of maps takes around 2 minutes per minute of video. The time needed for identification depends strongly on the number of animals, because the number of comparisons between maps grows with the square of the number of individuals. For 2 individuals identification takes around 2 minutes per minute of video, while for 10 individuals it may take up to 30 minutes per minute of video.

[19] Bought in 2010

In order to prevent saturation of the lab's computers, part of the process can be performed in a remote computer cluster.[20] This process is completely automated, so for the user there is almost no difference between doing the tracking locally or remotely. When option *Trueno* is selected, the program performs the segmentation locally[21] and then uploads the results to the cluster, sends the instruction to continue the rest of the process there and downloads the results once the tracking is finished. All communication between the local computer and the cluster is done using SSH.

[20] We use CSIC's cluster *Trueno*.

[21] It is better to perform this step in the local computer, because the videos are too heavy to upload to the cluster.

6.4 Set-up and conditions

General conditions

Because the method is based on the recognition of indviduals from their images, the main condition needed is enough quality of such images. Resolution must be high enough, meaning that the image of each individual must cover an area of at least approximately 100 pixels. Also, the video must be either uncompressed or compressed with HD high quality codecs. The camera must be of reasonable quality, but most modern regular cameras are adequate.[22]

[22] We use a regular video surveilance camera, see next section. We have also successfully tracked videos recorded with consumer video cameras.

The identification method is not scale-invariant, so the animals must be at roughly the same distance to the camera during the whole experiment. The system is probably robust to small changes in scale, but we have not tested depths larger 5% the distance between the

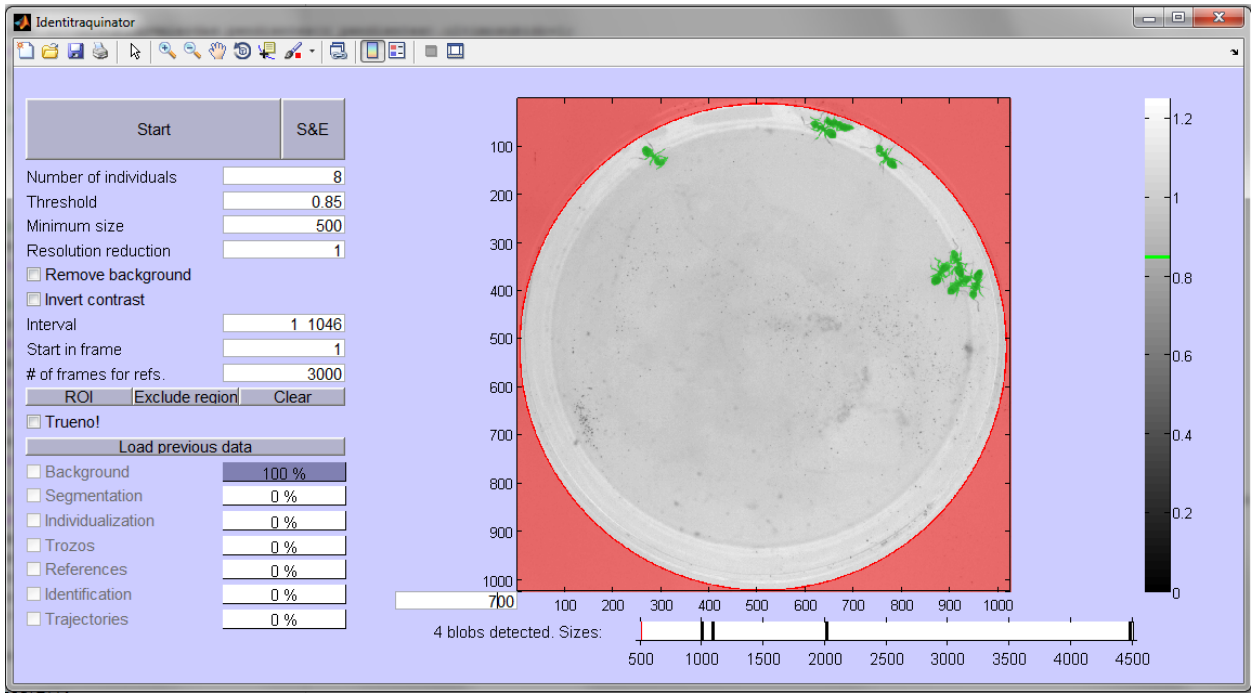


Figure 6.8: Interface of the tracking program.

camera and the set-up.

Illumination must be uniform in the whole set-up.

In order to be able to identify an individual its image must be segmented in a consistent fashion in all positions of the set-up. This is not typically the case with background subtraction routines, because usually more pixels of the animals are segmented when they are in regions where contrast with the background is higher. In order to have a more consistent segmentation, we do it by thresholding the intensity itself, and not the intensity minus the background. The method does incorporate a background removal routine that will eliminate any objects in the background, but those frames in which an individual is passing in front of a big background object are not segmented reliably enough for the identification, and must be discarded. Therefore, while the program can work with non-uniform background, its performance may be reduced.

These specifications are easy to implement, as evidenced by the fact that it works very well not only for videos recorded in our laboratory, but also for videos recorded by collaborators in Swansea University (UK) and the Centre de Recherches sur la Cognition Animale (Toulouse, France), in some cases in set-ups that were not designed following our specifications.

Description of the set-up

This section describes the set-ups used in our laboratory. The conditions described here must not be understood as necessary in all cases.

All videos recorded in our laboratory have been recorded with a

Basler A622f camera. This camera is monochrome and has a resolution of 1280 × 1024 pixels. It is connected to the computer via Firewire (IEEE 1394), and at full resolution has a frame rate of around 25 fps (the frame rate increases when we decrease resolution using only one part of the camera's sensor. We usually use only a square region in the middle of the sensor, so most of the videos are recorded at around 30 fps). The videos are directly recorded on the computer's hard disk, their length being limited only by the hard disk capacity (around 20 hours of uncompressed video in a 2 TB hard disk). The rest of the set-up changes for the different species.

FISH. For zebrafish (*D. rerio*), nine-spined stickleback (*P. pungitius*) and Medaka (*O. latipes*), videos were recorded in a 50 × 50 × 2 cm (length × width × height) arena, made from transparent perspex. This arena is completely closed (including a transparent roof), and fully submersed into a bigger tank (90 × 120 × 20 cm) equipped with a water acclimatizer to keep good conditions for the fish. The arena is sustained by four legs at around 5 cm of the white floor of the bigger tank, so that the shadows of the fish are diffused, facilitating the segmentation.[23] The camera is situated over the set-up at a distance of about 1.2 m, pointing directly downwards and equipped with an objective of 16 mm focal length, Pentax C31634KP - C1614-M (KP). In these conditions, the set-up covers around 950 × 950 pixels in the image. In order to have indirect and uniform illumination, we use 6 halogen floodlights (500 W each) pointing to the ceiling. A brown surface of 120 × 150 cm at the level of the camera prevents the light directly reflected on the ceiling to reach the set-up, so that illumination is indirect. Also, this brown surface projects a dim and uniform reflection on the water surface.

[23] This set-up was designed to maximize the quality of the videos, but we have found that the system works well with more conventional set-ups. In particular, the system is working very well with videos recorded in a set-up in the CRCA at Toulouse that has opaque white walls, opaque white floor and no roof.

FISH FROM THE SIDE. Zebrafish were recorded in a 25 × 3 × 25 cm (length × width × height) chamber inside a bigger tank made of glass. The camera was the same as for the previous case, and was situated at 1 m from the set-up, pointing horizontally towards the 25 × 25 cm face of the set-up. Illumination was the same as for the previous case. Best contrast was obtained in this case against black background, so we placed a black curtain over the rear side of the tank to create a dark background. A black curtain was also placed around the camera's objective, in order to obtain a dim and uniform reflection from the glass tank.

MICE (*M. musculus*) were recorded inside a transparent plastic cage of size 30 × 47 × 35 cm (length × width × height). The cage was open at the top, the walls being high enough to prevent the mice from escaping. The bottom of the cage was covered with sawdust in order to improve the comfort of the animals. Camera and illumination were identical as for the fish-from-above set-up, with the camera at 110 cm from the floor of the set-up.

INSECTS. Flies (*D. melanogaster*) and ants (unknown species, collected from a park in Madrid) were recorded in a circular arena of diameter 5.5 cm. The floor of the arena is made of transparent perspex. Walls and roof consisted of a Petri dish placed upside down. The inside of the Petri dish was coated with Fluon (Polytetrafluoroethylene, Sigma-Aldrich product Number 665800). Fluon is slippery for most insects, preventing them to climb to the walls and roof. With this disposition we could record walking flies with no need of cutting their wings. The camera was placed below the set-up, at 10 cm of the floor of the set-up. The objective used in the other conditions cannot focus at such a short distance, so this time we used a Pentax C31635KP - C1614-5M (KP) (focal length 16 mm). The insects were therefore seen through the transparent floor, against the white background of the petri dish covered with Fluon. The set-up was surrounded by white curtains, and illumination was provided by 5 hallogen floodlights (500 W each) outside the curtains, and pointing towards the inside.

6.5 Validation

The performance of the method varies depending on the species (some species are easier to identify than others), the number of individuals (in general the probability of wrong identification will be higher the more individuals we have) and the conditions of the set-up (shadows, occluded regions, etc.). While we have found that it works for all species that we have tested and in all conditions that are compatible with those described in section 6.4, in principle one should perform a validation for each new condition before doing definitive experiments.

We performed the validation by manually correcting one video and comparing the output of the automatic algorithm with the manually corrected trajectories. As described in section 6.2, the trajectories are first divided into one-individual segments, and then these segments are identified using all their frames. Therefore, the probability of wrong identifications depends strongly on the length of the segments, being higher for short segments than for longer ones. Also, because consecutive frames in moments when the animal is not moving have very similar information, distance traveled during the one-individual segment is a better indicator of the certainty of the identification than the number of frames. Therefore, we perform the analysis of probability of mistake separately for segments of each length.

In this analysis we only include one-individual segments that truly belong to one individual, according to the manually corrected trajectories. We chose to do so because errors are extremely infrequent in blobs composed by several overlapping individuals.[24] Therefore, including these cases increases significantly the proportion of "correct" cases, but this proportion is not reflected in the trajectories.

Therefore, "Correct" means that the segment was identified, and

[24] An error in a multiple blob occurs if the tracking confuses it with a one-individual blob and assigns an identity to it.

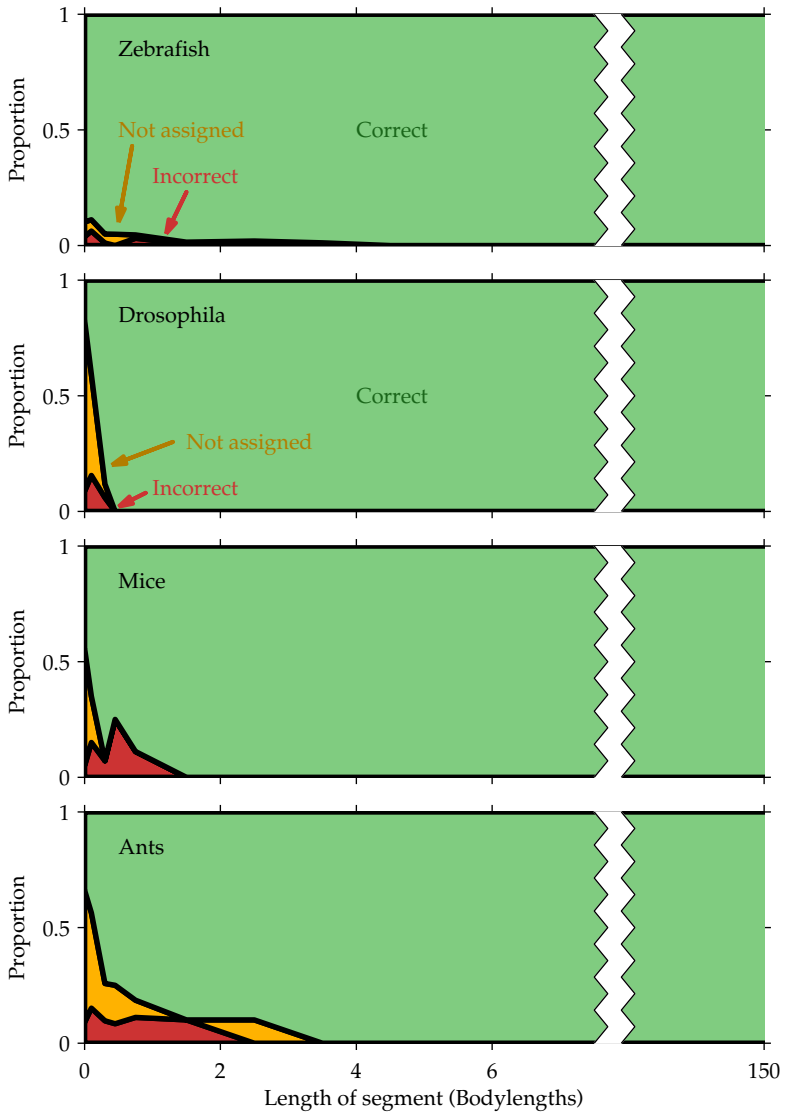


Figure 6.9: Proportion of one-individual segments of each length that have been assigned to the correct individual (green), not assigned to any individual (yellow) or assigned to the wrong individual (red). The length of one segment is defined as the distance traveled by the animal during the segment, and is scaled to the animal's body length to ease comparison between species. **a.** Results of a video of 8 zebrafish (*D. rerio*). **b.** Results of a video of 8 flies (*D. melanogaster*). **c.** Results of a video of 2 black mice (*M. musculus* **d.** Results of a video of 8 ants.

assigned to the correct individual. "Not assigned" means that the segment was not identified and therefore not included in the trajectories. "Incorrect" means that the segment was assigned, but to the wrong individual.

Figure 6.9 shows the probability of correct identification (green), ambiguous identification (orange) and wrong identification (red) as a function of the length of segments, and for different species. We find that only very short segments have a measurable probability of wrong identification. Because problematic segments are the shortest ones (that contribute to only a very small fraction of the trajectory), the final proportion of trajectories with mistakes is in all cases very low (<1% for all species).

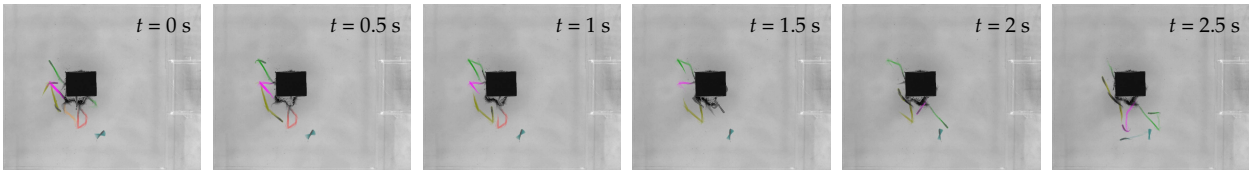


Figure 6.10: Zebrafish with a shelter, see complete video at youtu.be/kDlFo54MMrA The irregular gray silhouette around the black rectangle indicates that the tracking algorithm is also segmenting part of the rectangle's shadow. The program has no indication that the black rectangle (and its shadow) are a different object than the fish, and their blobs undergo the same process as those belonging to the fish. But these blobs are easily distinguished from the typical shape of an individual fish, and not taken into account for the trajectories.

6.6 Further advantages

Long videos

Because the reliability of the identities does not depend on the length of the video, and because the system works with conventional cameras, there is no upper limit on the duration of the videos.

Relation between videos of different days

For many experiments it is necessary to record several different videos of the same individuals, sometimes even in different days (for example, different trials in a learning experiment). Our system allows, not only to obtain the trajectory of each individual in each video, but also to relate the results of the different videos, so that it is possible to monitor the behavior of each individual in all videos.

There are two different ways to analyze these kind of videos. One may use the references acquired in one of the videos to perform the tracking in all of them. This method has the advantage that only one video needs to be long enough to acquire references, the rest being arbitrarily short. Another alternative is to perform the tracking of each video individually, acquiring a set of references for each of them, and afterwards compare the references of the different videos.

It is possible that the aspect of animals changes over time, and if these changes are big the algorithm will not be able to re-identify the individuals. This will be the case, for example, with videos of zebrafish larvae recorded in different days, because the larvae grow significantly from one day to the next. Adults may also change their appearance, but we have observed that we can routinely re-identify the same individuals in videos of different days. For example, we have performed experiments with the same group of zebrafish every working day for two weeks. We have also successfully related videos of Medaka that were recorded one week apart.

Animals can go out of view

Our method does not require all animals to be visible during the whole video, because an animal that goes out of view will be re-identified when it comes back. Therefore, it is valid to perform tracking in experiments in which the camera only covers part of the experimental arena, or in which there are shelters (for example as in Figure 6.10).

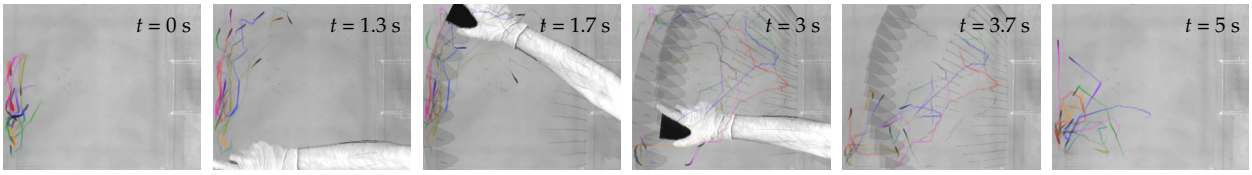


Figure 6.11: Illustration of the robustness of the algorithm to severe disturbances. The experimenter's hand is holding a black object that passes the segmentation threshold, but this object is not confused with the fish (notice that the trail of silhouettes produced by the object is gray, denoting that they were not identified as any of the individuals). The system is robust both to the hand's occlusions and to the extraneous object, providing reliable trajectories before and after the disturbance.

Identification

The method can also be used to identify individuals, for example to sort them after an experiment in which they are mixed. For this purpose, previous references of each individual must be acquired (either by recording each individual separately, or by recording a video with several individuals of the same type together). Once the references are acquired the animals can be mixed and the experiments performed. After the experiment, one must record a short video of each individual, in order to re-identify it. A few seconds of video usually suffice, because it is enough to wait until the animal has moved 10-20 bodylengths to have a very reliable identification (Figure 6.9).

Robustness

A perturbation at one point of the video may alter the results in several ways. First, all or some of the identities may be lost during the disturbance. This is not a problem, because the program does not need to have a continuous and unbroken set of trajectories. The identities may not be reliable during the disturbance, but they will be unaffected before and after the disturbance (Figure 6.11).

More important problems would arise if the disturbances affected the process of collection of references. But again this process does not require any continuity in the video, the references being collected from short disconnected periods along the whole video. And because of the strict conditions that must be fulfilled for an image to be included in the references, we have never observed that they are contaminated by moments in which a disturbance is present. Therefore, it is not necessary to manually exclude the disturbed period from the tracking, the algorithm being robust enough to deal with it.

Bibliography

Attanasi, A., Cavagna, A., Del Castello, L., Giardina, I., Jelic, A., Melillo, S., Parisi, L., Shen, E., Silvestri, E., and Viale, M. (2013). Tracking in three dimensions via multi-path branching. *ArXiv*, 1305.1495v.

Balcells, M., DeMenthon, D., and Doermann, D. (2005). An appearance-based approach for consistent labeling of humans and objects in video. *Pattern Analysis and Applications*, 7(4):373–385.

Branson, K. and Belongie, S. (2005). Tracking multiple mouse contours (without too many samples). *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:1039 – 1046.

Branson, K., Robie, A. A., Bender, J., Perona, P., and MH, D. (2009). High-throughput ethomics in large groups of Drosophila. *Nature Methods*, 6(6):451–457.

Butail, S. and Paley, D. A. (2012). Three-dimensional reconstruction of the fast-start swimming kinematics of densely schooling fish. *Journal of the Royal Society, Interface*, 9(66):77–88.

Dahlbom, S. J., Lagman, D., Lundstedt-Enkel, K., Sundström, L. F., and Winberg, S. (2011). Boldness predicts social status in zebrafish (Danio rerio). *PloS one*, 6(8):e23565.

de Chaumont, F., Coura, R. D.-S., Serreau, P., Cressant, A., Chabout, J., Granon, S., and Olivo-Marin, J.-C. (2012). Computerized video analysis of social interactions in mice. *Nature methods*, 9(4):410–417.

Delcourt, J., Becco, C., Vandewalle, N., and Poncin, P. (2009). A video multitracking system for quantification of individual behavior in a large fish shoal: advantages and limits. *Behavior research methods*, 41(1):228–235.

Delcourt, J., Denoël, M., Ylieff, M., and Poncin, P. (2012). Video multitracking of fish behaviour: a synthesis and future perspectives. *Fish and Fisheries*, 14(2):186–204.

Delcourt, J., Ylieff, M., Bolliet, V., Poncin, P., and Bardonnet, A. (2011). Video tracking in the extreme: a new possibility for tracking nocturnal underwater transparent animals with fluorescent elastomer tags. *Behavior research methods*, 43(2):590–600.

Dennis, R. L., Newberry, R. C., Cheng, H.-W., and Estevez, I. (2008). Appearance matters: artificial marking alters aggression and stress. *Poultry science*, 87(10):1939–1946.

Flusser, J. and Suk, T. (1993). Pattern recognition by affine moment invariants. *Pattern recognition*, 26(1):167–174.

Fontaine, E., Lentink, D., Kranenbarg, S., Müller, U. K., van Leeuwen, J. L., Barr, A. H., and Burdick, J. W. (2008). Automated visual tracking for studying the ontogeny of zebrafish swimming. *The Journal of experimental biology*, 211(Pt 8):1305–1316.

Gabriel, P., Verly, J., Piater, J., and Genon, A. (2003). The state of the art in multiple object tracking under occlusion in video sequences. *Advanced Concepts for Intelligent Vision Systems*, pages 166–173.

Herbert-Read, J., Perna, A., Mann, R., Schaerf, T., Sumpter, D., and Ward, A. (2011). Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences USA*, 108(45):18726–18731.

Huang, J., Ravi Kumar, S., Mitra, M., Zhu, W., and Zabih, R. (1999). Spatial color indexing and applications. *International Journal of Computer Vision*, 35(3):245–268.

Jerry, D. R., Stewart, T., Purvis, I. W., and Piper, L. R. (2001). Evaluation of visual implant elastomer and alphanumeric internal tags as a method to identify juveniles of the freshwater crayfish, Cherax destructor. *Aquaculture*, 193(1-2):149–154.

Kato, S., Nakagawa, T., Ohkawa, M., Muramoto, K., Oyama, O., Watanabe, A., Nakashima, H., Nemoto, T., and Sugitani, K. (2004). A computer image processing system for quantification of zebrafish behavior. *Journal of neuroscience methods*, 134(1):1–7.

Khan, Z., Balch, T., and Dellaert, F. (2005). MCMC-based particle filtering for tracking a variable number of interacting targets. *IEEE transactions on pattern analysis and machine intelligence*, 27(11):1805–1819.

King, A. J., Wilson, A. M., Wilshin, S. D., Lowe, J., Haddadi, H., Hailes, S., and Morton, A. J. (2012). Selfish-herd behaviour of sheep under threat. *Current biology*, 22(14):R561–R562.

Kovalev, V. and Petrou, M. (1996). Multidimensional co-occurrence matrices for object recognition and matching. *Graphical Models and Image Processing*, 58(3):187–197.

Lewejohann, L., Hoppmann, A. M., Kegel, P., Kritzler, M., Krüger, A., and Sachser, N. (2009). Behavioral phenotyping of a murine model of Alzheimer's disease in a seminaturalistic environment using RFID tracking. *Behavior research methods*, 41(3):850–856.

Lowe, D. G. (1999). Object recognition from local scale-invariant features. *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 2:1150–1157.

McKenna, S. J., Jabri, S., Duric, Z., Rosenfeld, A., Wechsler, H., and A, R. (2000). Tracking groups of people. *Computer Vision and Image Understanding*, 80(1):42–56.

Molet, M., Chittka, L., Stelzer, R. J., Streit, S., and Raine, N. E. (2008). Colony nutritional status modulates worker responses to foraging recruitment pheromone in the bumblebee Bombus terrestris. *Behavioral Ecology and Sociobiology*, 62(12):1919–1926.

Nagy, M., Akos, Z., Biro, D., and Vicsek, T. (2010). Hierarchical group dynamics in pigeon flocks. *Nature*, 464(7290):890–893.

Okuma, K., Taleghani, A., de Freitas, N., Little, J. J., and Lowe, D. G. (2004). A boosted particle filter: Multitarget detection and tracking. *Lecture Notes in Computer Science*, 3021:28–39.

Pentland, A., Moghaddam, B., and Starner, T. (1994). View-based and modular eigenspaces for face recognition. *Proceedings of the IEEE Conference on Vision and Pattern Recognition*, pages 84–91.

Robinson, E. J. H., Richardson, T. O., Sendova-Franks, A. B., Feinerman, O., and Franks, N. R. (2008). Radio tagging reveals the roles of corpulence, experience and social information in ant decision making. *Behavioral Ecology and Sociobiology*, 63(5):627–636.

Sanchez, O. and Dibos, F. (2004). Displacement Following of Hidden Objects in a Video Sequence. *International Journal of Computer Vision*, 57(2):91–105.

Sanjeev Arulampalam, M., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188.

Streit, S., Bock, F., Pirk, C. W. W., and Tautz, J. (2003). Automatic life-long monitoring of individual insect behaviour now possible. *Zoology (Jena, Germany)*, 106(3):169–71.

Takala, V. and Pietikainen, M. (2007). Multi-object tracking using color, texture and motion. *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–7.

Turk, M. and Pentland, A. (1991). Face recognition using eigenfaces. *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition*, pages 586–591.

Wiskott, L., Fellous, J.-M., Kuiger, N., and von der Malsburg, C. (1997). Face recognition by elastic bunch graph matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779.

Conclusions

– So what do you think?
– Well, whatever.

El disputado voto del señor Cayo
Miguel Delibes

- In the nervous system of *C. elegans*, the sensory-motor subnetwork is significantly nearer to its optimal configuration than the subnetwork of interneurons. While the macroscopic organization of the nervous system in 10 ganglia cannot be explained by wiring economy applied on the whole system, it emerges from its application to the sensory-motor subnetwork alone.

- We propose that deviations from optimality in many-component biological systems should follow a given pattern, deviations being larger for those components with smaller impact on the global fitness. The study of this pattern of deviations gives additional information about the shape of the fitness around the optimum.

- Deviations from the minimum wire configuration in the nervous system of *C. elegans* follow the expected pattern. The structure of these deviations indicates that wiring cost grows sublinearly with wiring length.

- Deviations from the maximum growth configuration in the metabolic network of *E. coli* follow the expected pattern. The structure of these deviations indicates that the network is optimized for biomass production and not for exclusive production of ATP.

- We have built a model of collective decision-making based on Bayesian integration of social and non-social information.

- The Bayesian model predicts three qualitative regimes, each of them characterized by the way in which animals count. Specifically, the model predicts that in some situations it is not necessary to count beyond a certain number. In other situations the model predicts the use of Weber's law to count high numbers but now for low numbers. In the third regime, the only relevant information comes from the difference of the two numbers, and not from their absolute values.

- The Bayesian model describes with high accuracy our experimental results in zebrafish (*D. rerio*) and previously published experimental data in three-spined stickleback (*G. aculeatus*) and argentine ants (*L. humile*).

- We have developed a tracking and identification technique that allows to follow the trajectory of each individual of a group during

experiments of arbitrary length. The technique is valid for many species, including *D. rerio*, *O latipes*, *G. aculeatus*, *M. musculus* and *D. melanogaster*.

Conclusiones

– Y usted ¿Qué dice?
– Que bueno.

El disputado voto del señor Cayo
Miguel Delibes

- En el sistema nervioso de *C. elegans*, la subred sensorimotora está significativamente más cerca de su configuración óptima que la subred de interneuronas. Mientras que la organizacion macroscópica del sistema nervioso (separación en 10 ganglios) no puede explicarse aplicando el principio de economía de cable sobre el sistema completo, emerge al aplicarlo sobre la subred sensorimotora exclusivamente.

- Postulamos que las desviaciones respecto del óptimo en sistemas biológicos con muchos componentes siguen un determinado patrón, con mayores desviaciones en las componentes que menos impacto tienen en la aptitud (*fitness*) total. El estudio de este patrón de desviaciones aporta infocmación adicional acerca de la forma de la función objetivo en las inmediaciones del óptimo.

- Las desviaciones respecto de la configuración de mínimo cableado en el sistema nervioso de *C. elegans* muestran el patrón esperado. La estructura de estas desviaciones indican que el coste de las conexiones crece de manera sub-lineal con su longitud.

- Las desviaciones respecto de la configuración que maximiza el crecimiento en la red metabólica de *E. coli* se corresponden con el patron esperado. La estructura de estas desviaciones indican que la red está optimizada para la produccion de biomasa, y no para la producción exclusiva de ATP.

- Hemos construido un modelo de toma de decisiones colectiva basado en la integración Bayesiana de informacion social y no-social.

- El modelo Bayesiano predice tres regímenes cualitativos, cada uno de los cuales está caracterizado por una manera en la que los animales deben contar. En concreto, el modelo predice que en algunas situaciones no es necesario contar a partir de un cierto número. En otras situaciones el modelo predice la ley de Weber para contar números grandes, pero no para números pequeños. En otros casos, la única informacion relevante es la diferencia entre dos cantidades, y no el valor absoluto de cada una.

- El modelo Bayesiano describe con gran precisión nuestros resultados experimentales en pez cebra (*D. rerio*) y resultados exper-

imentales previamente publicados en espinosillo (*G. aculeatus*) y hormiga argentina (*L. humile*).

- Hemos desarrollado una técnica de seguimiento e identificación, que permite seguir la trayectoria de cada individuo de un grupo durante experimentos arbitrariamente largos. La técnica es válida para muchas especies, incluyendo *D. rerio*, *O latipes*, *G. aculeatus*, *M. musculus* y *D. melanogaster*.